Tampere University

ATTE NIKKILÄ

Low doses of ionizing radiation and the risk of childhood leukemia

Tampere University Dissertations 122

Tampere University Dissertations 122

ATTE NIKKILÄ

Low doses of ionizing radiation and the risk of childhood leukemia

ACADEMIC DISSERTATION
To be presented, with the permission of
the Faculty Council of Social Sciences
of Tampere University,
for public discussion in the auditorium F211A+B
of the Arvo building, Arvo Ylpön katu 34, Tampere,
on 27.9.2019, at 12 o'clock.

ACADEMIC DISSERTATION
Tampere University, Faculty of Social Sciences
Finland

Responsible supervisor and Custos	Prof. Anssi Auvinen Tampere University Finland	
Supervisor	Adj. Prof. Olli Lohi Tampere University Finland	
Pre-examiners	Prof. Arja Harila-Saari Karolinska Institutet Sweden	Adj. Prof. Laura Madanat-Harjuoja Harvard Medical School USA
Opponent	Prof. Mark Pearce Newcastle University United Kingdom	

The originality of this thesis has been checked using the Turnitin OriginalityCheck service.

Cover design: Roihu Inc.

ISBN 978-952-03-1228-2 (print)
ISBN 978-952-03-1229-9 (pdf)
ISSN 2489-9860 (print)
ISSN 2490-0028 (pdf)
http://urn.fi/URN:ISBN:978-952-03-1229-9

PunaMusta Oy – Yliopistopaino
Tampere 2019

To Laura.

For her understanding, and her endless support.

ACKNOWLEDGEMENTS

The studies presented in this dissertation were carried out at the Tampere Center for Child Health Research of Tampere University between 2014 and 2019.

I am extremely grateful to both of my supervisors. Professor Anssi Auvinen and adjunct professor Olli Lohi have led me to a life-long path of science. From the very beginning I have been fascinated by their incredible work ethic and passion for advancing science in their own fields. I have had the feeling that I was allowed to work with researchers from the highest tiers of scientific work and at times I wondered whether their time would be better spent with more promising doctoral candidates. However, their positive feedback and words of encouragement had a tremendous effect on me. My feelings might have sailed at times, but never did I feel anything but motivated after our group meetings.

Anssi, working with you taught me how to view world through scientist's spectacles. Critical and fair assessment of published work, including my own work, is crucial for the scientific process. Now I feel I am able to achieve this using the methods you have taught me, and it will bear fruit throughout my career. Olli, observing the way you are able to balance clinical work in a leading position alongside managing your own top-level research group has shown me that what I once would have thought impossible, can be achieved with hard work, dedication and talent. I am forever grateful for your continuous support and for the numerous doors you have opened for me.

The early and intensive support from our group's statistician Jani Raitanen was essential. It allowed me to begin exploring the software and coding-oriented sea of epidemiology. Soon, I was able to swim on my own but even after that, our numerous discussions after working hours helped me forward during the times of desperation.

In a way, I have had a third supervisor. Doctor Hannu Arvela's input regarding the exposure assessment and data from Radiation and Nuclear Safety Authority has been irreplaceable. To iterate, completing this thesis would not have been possible without him. I thank him also for his patience with our lengthy email discussions and occasional phone calls.

I would also like to thank all of my other co-authors (Sini Erme, Olli Holmgren, Gerald Kendall, Ben Spycher, Juha Mehtonen and Merja Heinäniemi). Their expertise and contributions made this dissertation happen. I am especially thankful to professor Gerald Kendall from University of Oxford for orchestrating our study on the effects of complete residential histories. I have received numerous helpful comments from other experts in our field and I would like to thank everyone behind them.

I am also grateful to the members of my follow-up group. Professors Per Ashorn and Nea Malila were always swift to answer in the issues regarding the progress of my doctoral studies. I am thankful for the insightful comments of the official reviewers of this dissertation. The suggestions from professor Arja Harila-Saari and associate professor Laura Madanat-Harjuoja allowed this thesis to reach its full potential.

The support from my family has been hugely important to me and for that I will thank my two sisters, dad and mom. Regardless of everything written before, the single force pushing me through the obstacles was the never-ending support from my wife, Laura, who I love so much.

ABSTRACT

Despite the extensive scientific work in the field of childhood leukemia, the etiology of the disease has remained elusive. High-dose ionizing radiation is a well-established risk factor of childhood leukemia, however, marked uncertainty prevails regarding the effect of low doses. Studies exploring the association of leukemia with background radiation, computerized tomography scans and high indoor radon concentrations have been published but the results have been conflicting, and clinically and etiologically interesting subgroups have not been always reported. Studying the subject using high-quality, nationwide, population and register-based Finnish data with comprehensive information on exposures, including complete residential histories, provide a unique opportunity to advance the field.

The aim of this study was to evaluate the association between childhood leukemia and exposure to background radiation, computerized tomography scans and indoor radon using a case-control design in a nationwide register setting. The red bone marrow dose estimation was carried out using the best available prediction methods and their performance was evaluated separately. A separate statistical model was developed to predict indoor radon concentrations. In the risk analyses, clinically relevant subgroups were analyzed separately. The significance of complete residential histories, in studies of background radiation, was further investigated.

We identified all childhood leukemia cases diagnosed in Finland between 1990 and 2011 (N = 1093) from the Finnish Cancer Registry and individually matched them by gender and year of birth with thrice as many randomly sampled controls from the Population Register Centre. From the Radiation and Nuclear Safety Authority, we obtained 8 x 8 km square-maps of the natural background radiation and Chernobyl fallout, and direct measurements of indoor radon concentrations in over 80 000 residences. All electronically available pediatric computerized tomography scans were collected from the ten largest Finnish hospitals. Exposures were assessed using complete residential histories for each subject and the properties of the buildings were accounted for. A two-year latency-period was assumed. Information on birth weight, maternal smoking, Down syndrome, congenital malformations, parental socioeconomic status and parental education was obtained

to be included in the conditional logistic regression and exact conditional logistic regression models as potential confounders.

We observed an odds ratio of 1.01 (95% CI 0.97, 1.05) for every 10 nSv/h increase in average dose rate to red bone marrow from background radiation. The red bone marrow dose from pediatric computerized tomography scans showed a statistically significant increase per 1 mGy (Odds Ratio (OR) = 1.13, 95% CI 1.02, 1.26). We analyzed the quartiles of indoor radon concentration and observed no significant difference between groups. The odds ratios for the second, third and fourth quartiles were 1.08 (95% CI 0.77, 1.50), for the third 1.10 (95% CI 0.79, 1.53) and for the highest 1.29 (95% CI 0.93, 1.77), respectively.

In subgroup analyses of background radiation, the youngest age-group of 2–6.99 years showed a higher odds ratio of 1.27 (95% CI 1.01, 1.60) for every 1 mSv increase in cumulative dose to red bone marrow when compared to the age-group of 7–14.99 years. Moreover, results suggestive of a larger effect in cases with high hyperdiploidy were observed from natural background radiation. Slightly higher estimates were observed for young precursor B-cell acute lymphoblastic leukemia patients from computerized tomography scans and higher indoor radon concentrations.

The indoor radon concentration prediction model reached an r^2 of 0.21 for houses and 0.20 for apartments using a log-linear approach. Random forest models performed better reaching a coefficient of determination of 0.28 for houses and 0.23 for apartments. Neither approach showed signs of considerable overfitting and their performance was adequate in internal validation.

In analyses of residential histories, the median distance between subject's successive dwellings was 3.4 km and the median difference in the indoor red bone marrow dose rate from background radiation was 2.9 nSv/h. The dose rates in successive dwellings were correlated (0.62, 95% CI 0.60, 0.64, Pearson). There was a high concordance between odds ratios for average dose rate to red bone marrow when evaluated based on only the first dwelling (OR = 1.02, 95% CI 0.99, 1.05), the last dwelling (OR = 1.00, 95% CI 0.98, 1.03) and for a subgroup of subjects who had lived in only one dwelling (OR = 1.05, 95% CI 0.98, 1.10).

In general, the results provide support to the notion that even low doses of ionizing radiation are harmful to children and increase the risk for childhood leukemia. In exploratory subgroup analyses, higher estimates were observed for younger age group at diagnosis, precursor B-cell acute lymphoblastic leukemia and the genetic subgroup with high hyperdiploidy but these results require confirmation in an independent dataset. When available, full residential histories should be

prioritized. Nevertheless, informative results can be produced using information limited to the address of birth or the address of diagnosis. In contrast to the other studies, the risk estimates for computerized tomography scans were higher than expected based on Japanese atomic bomb studies and previous publications. This is probably due to a degree of random error and the potential effect of unknown predisposing factors cannot be excluded. Also, predictive modelling of indoor radon exposure carries major uncertainties and the results need to be interpreted with caution.

TIIVISTELMÄ

Vuosikymmeniä jatkuneesta laadukkaasta tutkimustyöstä huolimatta lasten leukemian syytekijät ovat yhä suurelta osin hämärän peitossa. Suuriannoksinen ionisoiva säteily on yksi lasten leukemian hyvin tunnetuista riskitekijöistä, mutta pienten annosten merkitys on epäselvä. Luonnollista taustasäteilyä, lasten tietokonetomografioita ja korkeita sisäilman radonpitoisuuksia on tutkittu lasten leukemian riskitekijöinä, mutta tutkimusten tulokset ovat olleet ristiriitaisia ja niissä on harvoin pystytty tarkastelemaan kaikkia kliinisesti kiinnostavia alaryhmiä erikseen. Niinpä aiheen tutkiminen on tarpeellista hyödyntäen laadukasta suomalaista rekisteripohjaista dataa, jossa on mukana myös kattava asuinhistoriatieto.

Tämän väitöstutkimuksen tavoitteena oli arvioida lasten leukemiariskiä luonnollisesta taustasäteilystä, tietokonetomografioista ja korkeista sisäilman radonpitoisuuksista käyttäen koko maan laajuista rekisteripohjaista tapaus-verrokkiasetelmaa. Punaisen luuytimen annosten arviot suoritettiin parhailla tunnetuilla menetelmillä ja niiden onnistumista arvioitiin erikseen. Sisäilman radonpitoisuuksien estimointiin kehitettiin tilastollinen ennustusmalli. Riskianalyyseissä kliinisesti ja etiologisesti kiinnostavat alaryhmät arvioitiin erikseen. Kattavien asuinhistorioiden merkitystä arvioitiin erikseen luonnollisen taustasäteilyn kontekstissa.

Keräsimme kaikki alle 15-vuotiaana lasten leukemiaan diagnosoidut suomalaiset lapset ajalta 1990–2011 (N = 1093) Syöpärekisteristä. Heille valittiin satunnaisotannalla Väestörekisterikeskuksesta kolminkertainen määrä sukupuolen ja syntymävuoden mukaan yksilökaltaistettuja verrokkeja. Säteilyturvakeskukselta saatiin 8 x 8 km neliökartat maaperän luonnollisesta taustasäteilystä sekä Tšernobylin ydinonnettomuuden laskeumasta. Säteilyturvakeskus luovutti lisäksi yli 80 000 suomalaisissa asunnoissa tehtyä sisäilman radonpitoisuuden mittausta. Kaikki sähköisesti saatavilla olleet lasten tietokonetomografiatutkimukset kerättiin kymmenestä suurimmasta sairaalasta. Säteilyaltistuminen arvioitiin hyödyntäen kattavia asuinhistorioita samalla huomioiden sen suuruuteen vaikuttavat tekijät. Analyyseissä käytettiin kahden vuoden latenssiaikaa. Regressiomallien vakiointia varten kerättiin tieto seuraavista mahdollisista sekoittavista tekijöistä: lapsen syntymäpaino, äidin raskauden aikainen tupakointi, Downin syndrooma,

synnynnäiset epämuodostumat, vanhempien sosioekonominen asema, vanhempien koulutustaso. Tilastolliset analyysit toteutettiin käyttäen ehdollista logistista regressiota ja eksaktia ehdollista logistista regressiota.

Havaitsimme maaperän taustasäteilyyn liittyvän lasten leukemian ristitulosuhteen marginaalisen kohoamisen jokaista punaisen luuytimen annosnopeuden 10 nSv/h nousua kohti (OR = 1,01; 95 % LV 0,97; 1,05). Lasten tietokonetomografioista peräisin olevaan punaisen luuytimen säteilyannokseen liittyi kohonnut lasten leukemian riski (OR = 1,13; 95 % LV 1,02; 1,26, jokaista 1 mGy annosta kohti). Vertasimme sisäilman radonaltistumisen neljänneksiä toisiinsa ja emme havainneet tilastollisesti merkitseviä eroja ryhmien välillä. Ristitulosuhde toiselle neljännekselle oli 1,08 (95 % LV 0,77; 1.50), kolmannelle 1,10 (95 % LV 0,79; 1,53) ja neljännelle 1,29 (95 % LV 0,93; 1,77).

Alaryhmäanalyyseissä havaitsimme taustasäteilyn vaikutuksen olevan merkitsevästi suurempi alle 7-vuotiaiden ikäryhmässä (OR = 1,27; 95 % LV 1,01; 1,60 jokaista 1 mSv punaisen luuytimen annoksen nousua kohti). Havaitsimme myös viitteellisiä tuloksia korkeammasta riskistä tiettyyn geneettiseen alaryhmään (hyperdiploidi) kuuluville potilaille. Lisäksi tietokonetomografioiden ja sisäilman radonpitoisuuden aiheuttamat altistukset aiheuttivat suurempia riskejä nuorimpien lasten akuuttiin B-soluleukemiaan liittyen mutta nämä tulokset vaativat varmistuksen itsenäisellä aineistolla.

Sisäilman radonpitoisuuksien ennustusmalli saavutti 0,21 selitysasteen pientaloille ja 0,20 selitysasteen kerrostaloille log-lineaarisella mallilla. Satunnaismetsä-mallit toimivat paremmin ja saavuttivat 0,28 selitysasteen pientaloille ja 0,23 selitysasteen kerrostaloasunnoille. Kumpaankaan lähestymistapaan ei liittynyt voimakasta ylisovittamista ja omalla datalla malleja validoitaessa mallit olivat vakaita.

Mediaanietäisyys tutkimushenkilöiden kahden peräkkäisen asunnon välillä oli 3,4 km ja ero asunnon sisätilojen annosnopeuksissa oli 2,9 nSv/h. Perättäisten asuntojen annosnopeudet korreloivat keskenään (0,62; 95% CI 0,60; 0,64; Pearson). Taustasäteilyperäisen punaisen luuytimen saaman säteilyn annosnopeuteen liittyvät ristitulosuhteen säilyivät saman suuntaisina tarkasteltaessa ainoastaan tutkimushenkilöiden syntymäasuntoa (OR = 1,02; 95 % LV 0,99; 1,05) tai diagnoosiasuntoa (OR = 1,00; 95 % LV 0,98; 1,03) mutta hieman suurempi ristitulosuhde havaittiin analysoitaessa erikseen alaryhmää, jonka jäsenet eivät koko asuinhistoriansa aikana olleet asuneet kuin yhdessä asunnossa (OR = 1,05; 95 % LV 0,98; 1,10).

Tuloksemme tukevat käsitystä, että pienetkin ionisoivan säteilyn annokset ovat haitallisia lapsille ja nostavat lasten leukemiariskiä. Alaryhmäanalyyseissä havaittiin kohonneita riskisuureita nuoremmille akuuttia B-soluleukemiaa sairastaville sekä hyperdiploidista leukemiamuotoa sairastaville, mutta nämä tulokset vaativat vahvistuksen itsenäisellä aineistolla. Tulostemme mukaan kattavaa tietoa asuinhistorioista on suosittava, kun ne ovat saatavilla. Hyödyllisiä riskiarvioita pystytään kuitenkin muodostamaan käyttäen aineistoja, joissa osoitetiedot rajoittuvat diagnoosi- tai syntymäosoitteeseen. Eroten muista osajulkaisuista tietokonetomografiaperäisen säteilyannoksen riskisuureet olivat suurempia kuin Japanin atomipommitutkimuksiin ja aiempiin julkaisuihin perustuen oli odotettavissa. Tämä aiheutui luultavasti osittain satunnaisvirheestä, mutta tuntemattomien sekä tietokonetomografialle että lasten leukemialle altistavien tekijöiden olemassaoloa ei pystytty täysin poissulkemaan. Lisäksi sisäilman radonpitoisuuksien mallintamiseen liittyi epävarmuustekijöitä ja mallin avulla muodostettuja riskiarvioita tulee tulkita varoen.

CONTENTS

List of Figures

List of Tables

ABBREVIATIONS

ALARA	As low as reasonably achievable
ALL	Acute lymphoblastic leukemia
AML	Acute myeloid leukemia
CI	Confidence interval
CLL	Chronic lymphocytic leukemia
CML	Chronic myeloid leukemia
CLIC	Childhood Leukemia International Consortium
CT	Computerized tomography
DNA	Deoxyribonucleic acid
DS	Down syndrome
FRECCLE	Finnish Register-based Case-control study of Childhood leukemia
IARC	International Agency for Research on Cancer
ICRP	International Commission on Radiological Protection
IQR	Interquartile range
LNT	Linear no-threshold model
MI	Multiple imputation
MRI	Magnetic resonance imaging
NCICT	National Cancer Institute dosimetry system for Computed Tomography
OR	Odds ratio
PACS	Picture archiving and communications systems
pre-B-ALL	Precursor B-cell acute lymphoblastic leukemia
RNA	Ribonucleic acid
RR	Relative risk
RBM	Red bone marrow
SIR	Standardized incidence ratio
STUK	Radiation and Nuclear Safety Authority
T-ALL	T-cell acute lymphoblastic leukemia
UKCCS	United Kingdom Childhood Cancer Study

UNSCEAR	United Nations Scientific Committee on the Effects of Atomic Radiation

ORIGINAL PUBLICATIONS

I. **Nikkilä, A.,** Erme, S., Arvela, H., Holmgren, O., Raitanen, J., Lohi, O. & Auvinen, A. 2016. Background radiation and childhood leukemia: A nationwide register-based case-control study. *International Journal of Cancer*, vol. 139, no. 9, pp. 1975-1982.

II. **Nikkilä, A.***, Kendall, G.*, Raitanen, J., Spycher, B., Lohi, O. & Auvinen, A. 2018. Effects of incomplete residential histories on studies of environmental exposure with application to childhood leukaemia and background radiation. *Environmental Research*, vol. 166, pp. 466-472.

III. **Nikkilä, A.**, Raitanen, J., Lohi, O. & Auvinen, A. 2018. Radiation exposure from computerized tomography and risk of childhood leukemia: Finnish register-based case-control study of childhood leukemia (FRECCLE). *Haematologica*, vol. 103, no. 11, pp. 1873-1880.

IV. **Nikkilä, A.**, Arvela, H., Mehtonen, J., Raitanen, J., Heinäniemi, M., Lohi, O. & Auvinen, A. 2019. Predicting residential radon concentrations in Finland: Model development, validation, and application to childhood leukemia. *Submitted for publication.*

* Equal contributions

The permissions to re-use published articles were requested and obtained from the journals' offices. The original publications will be referenced later in the dissertation with Roman numerals (I-IV).

1 INTRODUCTION

1.1 Childhood leukemia

1.1.1 Definition and epidemiology

The first reported probable leukemia case was described in 1811 by Peter Cullen but the disease was better characterized by French surgeon Alfred Velpeau in 1826 (Piller, 2001). The first diagnosis was made by an Englishman John Bennett in 1845 (Piller, 2001). The term leukemia was introduced by Rudolf Virchow in 1845 and in 1868 Franz Neumann discovered the link of the disease with bone marrow (Piller, 2001). Farber reported the first antileukemic drug, aminopterin, in 1947 (Farber, Diamond, Mercer, Sylvester, & Wolff, 1948). At that time, the disease was fatal.

Leukemia is the most common malignant disease of childhood (Madanat-Harjuoja, Pokhrel, Kivivuori, & Saarinen-Pihkala, 2014; Steliarova-Foucher et al., 2017). Based on data from the Finnish Cancer Registry, the incidence of childhood leukemia in Finland has hovered around 50 new cases annually. Also, the incidence is slightly higher in boys as in other countries. The incidence in developed countries has been reported to be increasing slowly and overall a lower incidence has been reported in the African regions, which is likely partly due to underdiagnosis and non-systematic registration (Steliarova-Foucher et al., 2017). The typical incidence peak observed for ages 2 – 5 years is missing in developing countries suggesting the possibility of a contribution of etiological factors related to economic growth (Hrušák et al., 2002).

Due to intensive research efforts, the most modern treatment protocols already achieve a 5-year survival rate of over 90% depending on the ALL subtype (Toft et al., 2018). In Finland for years 2001 – 2010 the overall childhood leukemia 5-year survival rate was 83% (95% CI 76%, 88%) (Madanat-Harjuoja, Pokhrel, Kivivuori, & Saarinen-Pihkala, 2014). The most recent clinical research on ALL currently focuses on less common subtypes with poor prognosis and addresses the long- and short-term side-effects of given therapies by more accurate stratification strategies

and individualization of therapy (Toft et al., 2018). Simultaneously, the number of recognized distinct subtypes with targetable genetic lesions is increasing (Inaba, Greaves, & Mullighan, 2013).

The modern definition of leukemia covers a set of malignancies that originate from the blood-forming tissue and result in massive accumulation of undifferentiated and clonal blast cells in the bone marrow (World Health Organization, 2016). This inhibits the normal production and function of multiple components of the blood and bone marrow resulting in typical clinical manifestations and symptoms, namely hepatomegaly, splenomegaly, pallor, fever, bruising and pain (Clarke et al., 2016).

Leukemia consists of two major lineages, which both have acute and chronic forms, even though chronic leukemias are rare in children (Hutter, 2010). In fact, chronic lymphocytic leukemia (CLL) is not seen at all in children, i.e. the only chronic form seen in children is chronic myeloid leukemia (CML). Some cases also present with bi-lineage disease with attributes from both myeloid and lymphoid lineages. The most common subtype of leukemia in children is acute lymphoblastic leukemia (ALL) comprising 85% of all diagnoses. ALL can be further divided into T-cell and B-cell leukemias out of which B-cell leukemia, and more specifically precursor B-cell leukemia (pre-B-ALL), is the most common subtype. The other prevalent lineage is acute myeloid leukemia (AML) (Hutter, 2010). The different types of childhood leukemia have partly different risk factors and, for example, ionizing radiation has been shown to be associated with a higher risk of AML (Hsu et al., 2013).

Modern diagnostics employ multiple approaches, which aim at more personalized therapies based on risk-grouping mainly on the basis of treatment response and genetic subtype. Modern treatment protocols recognize adverse effects that the used therapies can lead to and aim at their minimization. At diagnosis, bone marrow samples are collected, and cerebrospinal fluid is analyzed to rule out possible central nervous system involvement. The morphology of the bone marrow is observed with light microscopy and the surface antigens are investigated with flow cytometry. Karyotyping or SNP profiling can reveal alterations in the number of chromosomes. Clones with numbers typically exceeding 50 and DNA-indices (deoxyribonucleic acid-index) above 1.16 are called high-hyperdiploid and cells with the number below 44 or DNA-indices below 0.85 are called hypodiploid (Paulsson & Johansson, 2009; Safavi et al., 2013). Specific genetic alterations can be tested for with fluorescent in-situ hybridization or PCR (e.g. *TEL-AML1* translocation) and can be indicators used to risk stratify patients. Next-generation sequencing

technologies are becoming more common as they reveal multiple alterations simultaneously (Levy & Myers, 2016).

In Finland, the treatment is based on collaborative Nordic protocols by the Nordic Society of Paediatric Haematology and Oncology (NOPHO). Generally, the treatment of ALL consists of induction, consolidation, delayed intensification and maintenance phases. Allogenic stem cell transplantation is reserved for patients with poor treatment response. For example, the NOPHO ALL-2008 protocol stratified patients into three distinct risk groups: standard risk, intermediate risk and high-risk (Toft et al., 2018). Craniospinal radiation therapy has been used in the past as prophylaxis of central nervous system (CNS) involvement, but its adverse effects have outweighed the benefits and it is no longer used except for rare relapsed CNS cases. Modern treatment strategies are able to achieve a 5-year event-free survival exceeding 90% in the youngest age-groups (1-9 years) (Toft et al., 2018). This is a remarkable feat as the disease was universally fatal in the 1940s, and even a decade after that (Farber et al., 1948) (Figure 1).

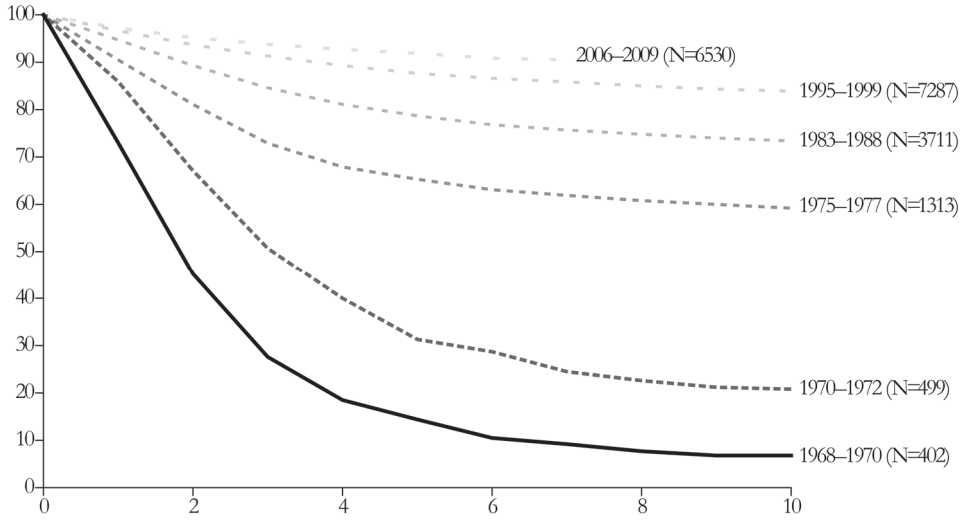


Figure 1. Childhood leukemia success story – Overall survival among children with ALL enrolled in Children's Cancer Group and Children's Oncology Group Clinical Trials 1968 – 2009. Adapted from Hunger and Mullighan. 2015.

1.1.2 Leukemogenesis

The modern understanding of developing overt leukemia suggests that two separate hits are needed (M. Greaves, 2018; Inaba et al., 2013). In this context a hit describes an event leading towards overt leukemia. The first predisposing hit happens *in utero* whereas the second occurs during childhood leading to overt disease. The understanding first came from studies on monozygotic twins and later from studying Guthrie cards (neonatal blood spots) and cord blood samples (Ford et al., 1993; Gale et al., 1997; M. Greaves, 2005; M. F. Greaves, Maia, Wiemels, & Ford, 2003; M. F. Greaves & Wiemels, 2003; Maia et al., 2004). The prevalence of certain predisposing hits has been debated in recent years but also proportions as high as 5% of the general population have been suggested in a common subtype (ETV6-RUNX1) of acute lymphoblastic leukemia (Schäfer et al., 2018). Also, on average childhood leukemia has a lower number of genetic alterations when compared to typical adult tumors (Pritchard-Jones, 1996). Due to the lower number of alterations, each of them is required to contribute more strongly to the hallmarks of cancer (Hanahan & Weinberg, 2011). Regardless of the increasing understanding of the disease, its etiology still remains relatively unknown (Eden, 2010).

There have been two major infection-related hypotheses for the development of childhood leukemia. Kinlen introduced a theory of population mixing as the major etiological component in 1988 (Kinlen, 2012). He proposed that an abnormal influx of people into an isolated rural site would introduce the native residents to an unknown burden of subclinical infections, which would then lead to leukemia after inappropriate immune response. Greaves' delayed infection hypothesis, first proposed in 1988, postulates that, in the absence of the expected load of infections in the early childhood, an exposure to common infectious diseases later in childhood might trigger leukemia as a rare response (M. Greaves, 2018). Greaves' hypothesis aimed to explain particularly the peak in ALL incidence in early childhood (2-5 years old). Neither of the theories suggests a specific pathogen.

1.1.3 Mortality and survival

Multiple advancements in treatment approaches of childhood leukemia have resulted in the excellent outcomes of today. The most recently completed Nordic treatment protocol of ALL reported a 5-year overall survival (OS) of 94% for young children of ages from one to nine (Toft et al., 2018). The respective event-free survival (EFS) was 89%. For children between 10 and 17 years of age the respective percentages

were 87% and 80%. For AML, the treatment outcomes are lower and in children overall survival rates of around 70% have been reported (Rubnitz, 2017). Regardless of the good progress, relapsed childhood leukemia is still the most common cause of death from cancer in children under 15 years of age in the United States (Nguyen et al., 2008).

Many clinical attributes of the disease affect the prognosis of the disease (Table 1). According to the latest Nordic protocol, attributes linked with poor response include higher age at diagnosis, higher white blood cell count at diagnosis, hypodiploidy, *KMT2A*-rearrengement and poor treatment response at early treatment. High hyperdiploidy and t(12;21)/ETV6-RUNX1 are associated with a better prognosis (Toft et al., 2018).

Table 1. A simplified summary of risk stratifying factors in the NOPHO ALL2008 treatment protocol (Toft et al. 2018)

Risk stratifying factor	Consequence
T-cell lineage	Dexamethasone instead of prednisolone in induction therapy
High leukocyte count (>100 10^9/l) at diagnosis	Dexamethasone instead of prednisolone in induction therapy
KMT2A gene rearrangement	High risk
Hypodiploidy	High risk
>25% blasts in bone marrow in dexamethasone induction arm (day 15)	High risk
Suboptimal (>0.1% blasts) therapy response at the end of induction (day 29)	Intermediate risk / High risk
Suboptimal (>0.1% blasts) response later in the treatment (day 79)	High risk
TCF3-PBX1 gene fusion	Intermediate risk / High risk
Intrachromosomal amplification of chromosome 21	Intermediate risk / High risk
Dicentric fusion chromosome from chromosomes 9 and 20	Intermediate risk / High risk
Central nervous system involvement after intrathecal therapy	Intermediate risk / High risk

1.2 Ionizing radiation

Ionizing radiation is defined as radiation with sufficient energy to detach electrons from the orbits of atoms or molecules. Losing an electron, in effect, gives the atom or molecule a negative charge thereby ionizing it and making it considerably more

reactive in addition to breaking its structure. (United Nations Scientific Committee on the Effects of Atomic Radiation, 2010)

1.2.1 Gamma radiation

Gamma radiation is a subtype of ionizing radiation. Electromagnetic radiation with extremely high frequency, or short wavelength ($<10^{-11}$ m), is defined as gamma radiation. The massless particles propagating the effect are called photons. Electromagnetic radiation with slightly longer wavelength ($<10^{-8}$ m) and lower energy is called X-rays. Gamma radiation is produced in nuclear reactions, decay of subatomic particles and in radioactive decay. Gamma radiation has high penetrance and thick concrete walls, or lead plates are used as protection. (International Agency for Research on Cancer, 2000)

In this dissertation the main source of terrestrial background gamma radiation are the natural decay chains of radioactive nuclei in the soil. The radium chain produces the majority of the terrestrial gamma radiation as it contains the Uranium-238 isotope, which is found abundantly in the Earth's crust in comparison to isotopes of other decay chains. Also, potassium-40 and thorium-232 contribute markedly to the terrestrial gamma radiation dose rate (International Agency for Research on Cancer, 2000).

A proportion of gamma radiation originates in space and is called cosmic radiation. The origin of cosmic radiation can be divided into two categories: solar and galactic. The average dose rate of cosmic radiation in Finland is 32 nSv/h and it is determined mostly by altitude from sea level. Also, higher dose rates of cosmic radiation are found in polar areas of the Earth (International Agency for Research on Cancer, 2000). The dose rates remain relatively similar everywhere in Finland due to the size of the country and the lack of variation in terrain elevation (Muikku, Bly, Kurttio, et al., 2014).

1.2.2 Alpha radiation

Alpha particles are nuclei of helium atoms. They consist of two neutrons and two protons effectively making them positively charged. Alpha radiation consists of alpha particles which are formed in a subclass of radioactive decay known as alpha decay. Radioactive decay is a process which allows an unstable atomic nucleus to stabilize by losing energy in the form of radiation. Contrary to gamma radiation, alpha-

radiation has extremely low penetrance and only a sheet of paper is enough to block the radiation. On the other hand, alpha nuclei are much more prone to releasing their energy into the material they interact with. (International Agency for Research on Cancer, 2000)

In the radium decay chain, radon-222 is a gaseous isotope, which can be inhaled and thus its and its daughter nuclei's decay can occur in the alveoli. Most of the deposited energy originates from two daughter nuclei of radon-222 and are called polonium-218 and polonium-214. The energy from inhaled radon is mainly deposited to the lungs and organ doses of other tissues are not as high. In total, inhaled radon-222 is the major source of alpha radiation in Finland (Muikku, Bly, Kurttio, et al., 2014).

1.2.3 Ionizing radiation as a carcinogen

In general, exposure to ionizing radiation has been associated with multiple adverse health effects. A uniform dose of 4 Gy will be lethal in 60 days to half of the adult population (ICRP, 2007). Deterministically high doses of ionizing radiation cause, in a time scale of hours to weeks, for example erythema and mucositis. These effects are known as tissue reactions. Late reactions from non-lethal doses occurring months from exposure include dermal necrosis, vascular occlusions, chronic infections and intestinal ulcerations. By contrast, these effects are known as stochastic effects of ionizing radiation. The severity and occurrence of symptoms is defined largely by the received radiation dose.

Stem cells of regenerating tissues are not able to maintain mitotic activity and the more differentiated cells are depleted. As a result, rapidly regenerating tissues suffer most. Deterministically ionizing radiation is also associated with syndromes including sterility, skin burns, hair loss, cataract, pneumonitis, bone marrow syndrome and gastro-intestinal syndrome. Stochastically, in addition to malignant diseases, there are results that show statistical evidence for heart disease, stroke, digestive disorders and respiratory diseases from lower doses of ionizing radiation. (ICRP, 2007)

In genetic analyses, the mutational fingerprint of ionizing radiation has been elusive to capture. There appear to be no specific sites where the damage occurs in the chromatin, however, it has been suggested that ionizing radiation could result in double strand break (DSB) clusters that could be differentiated from clusters resulting from other oxidative stress (Schipler et al., 2016). Ionizing radiation has

been referred to as 'the perfect carcinogen' as it has been found to promote every stage of cancer development. Though the association between high doses of ionizing radiation and leukemia is undisputed but the significance of low-dose ionizing radiation still remains unclear (Hsu et al., 2013).

In general, the carcinogenicity of ionizing radiation can be divided into direct and indirect harmful effects (International Agency for Research on Cancer, 2000). Indirect damage to DNA can occur by generation of free radicals, extremely reactive charged chemical compounds that are able to interact with intracellular molecules including chromatin causing damage to DNA and its supporting structures. Directly, ionizing radiation can also lead to double strand breaks, which can be repaired with a cellular process called non-homologous end-joining. Nevertheless, double strand breaks can result in translocations and ionizing radiation has been found to directly cause translocations (Reisz, Bansal, Qian, Zhao, & Furdui, 2014; Sigurdson et al., 2008). Translocations are a known marker of radiation exposure and Little et al. suggested that translocation-frequency analyses could offer independent estimates of exposure to ionizing radiation all the way down to doses of 100 mSv (Little et al., 2014). Also, bystander cells have been found to experience stress from ionizing radiation hitting neighboring cells and Azzam et al. report that mitochondrial DNA is also at risk and damages to it might explain delayed outcomes of ionizing radiation (Azzam, Jay-Gerin, & Pain, 2012).

1.2.4 Radiation dose

1.2.4.1 Principles of dose calculations

Radioactivity can be represented as the number of decays in a unit of time. The unit of activity is Becquerel (Bq = 1/s). The base of dose calculations is the absorbed dose, which is the amount of energy absorbed by a given mass. The unit of absorbed dose is Gray (Gy = J/kg). (International Agency for Research on Cancer, 2000)

A given absorbed dose has different biological effects determined by the type of radiation that carries the energy. The same absorbed dose from alpha radiation is, for example, twenty times as effective as the same dose from gamma radiation according to the International Commission on Radiological Protection (ICRP). When the biological effect of different radiation types is taken into consideration by a set of coefficients for different radiation types, a quantity called equivalent dose is formed. The unit of equivalent dose is the Sievert (Sv = J/kg). (ICRP, 2007)

In addition to the type of ionizing radiation, also the tissue receiving the energy bears significance. The sensitivity of different organs can be weighed with coefficients summing to one based on their radiosensitivity. For example, a given equivalent dose to red bone marrow (0.12) is weighted as twelve times more damaging than the same dose to the skin (0.01). A weighted average of different organ doses can be used to describe the total effect the radiation has on the individual. This quantity is called the effective dose and its unit is also Sievert (Sv = J/kg). (ICRP, 2007)

It is also necessary to define a quantity to represent the rate of dose accumulation. Dose rates are defined as the dose received during a given unit of time. The dose rates can be used similarly for all dose quantities (absorbed, equivalent and effective).

When describing exposure to residential radon concentrations, the quantity of interest is defined as a concentration. It represents the activity per unit of volume (Bq/m^3). This quantity is analogous to the dose rates. To describe cumulative exposure to residential radon, the concentration is integrated over time ($Bq/m^3 \cdot s$).

1.2.4.2 Dose to the red bone marrow

The main site of interest, regarding leukemogenesis, is the red bone marrow (RBM) and thus, in our studies, the quantity of interest is the organ dose to the RBM. The studied background radiation and radiation from computerized tomography scans (CT) consist of gamma radiation propagating quite freely through all tissues whereas the dose from inhaled radon is deposited primarily to lungs and the active bone marrow does not receive as much radiation from that source. However, regarding indoor radon, the dose to lymphatic cells in pulmonary tissue has also been suggested, by only few researchers, to be harmful (Harley & Robbins, 2009). The RBM doses from different types and sources of radiation have been modelled with human phantoms (Stovall et al., 2006). Multiple phantoms are required since the volume and the location of RBM, in both males and females, change as a function of age. Also, the amount, consistence and distribution of other shielding tissues change with age. In addition, the type of ionizing radiation (e.g. gamma vs. alpha) needs to be taken into consideration.

Computerized tomography is a means of medical imaging dependent on ionizing radiation. It uses a series of X-ray measurements from multiple angles to produce traditionally a series of cross-sectional representations of the target tissue. The RBM doses from pediatric computerized tomography scans have been modelled with 3D models and software-based approaches have proven to be accurate and the most

recent widely available software with sufficient capabilities is NCICT (Lee, Kim, Bolch, Moroz, & Folio, 2015). For estimation, the manufacturer and model of the scanner are required alongside other scanning parameters. In Finland, the scanning parameters are adjusted by hospital physicists locally to balance sufficient image quality and received dose. The fact that the scanning parameters are not standardized between different imaging sites poses a challenge in dose estimations. Also, nowadays CT scanners have specific pediatric settings to minimize exposure to ionizing radiation and prior to their existence, the doses to children were considerably higher (Kim et al., 2012). In recent years, approximately after year 2000, Finnish hospitals have started to record imaging studies electronically into picture archiving and communicating systems (PACS) and typically information on the doses is recorded as well.

Regarding RBM doses from terrestrial background radiation, the fact that indoor dose rates differ from outdoor dose rates at the same location needs to be taken into consideration. The indoor dose rates are affected by the shielding of the building materials but, on the other hand, the building materials themselves emit gamma radiation if they consist of rock-based materials (Arvela, Hyvönen, Lemmelä, & Castrén, 1995). Rock-based building materials are built from raw materials from the same area and, thus, the dose rate in the area would represent the dose rate emitted from the rock-based building materials in the area. Also, the occupancy, proportion of time spent indoors, needs to be modelled as the dose rates are different indoors and outdoors. Worldwide estimates are reported by the United Nations Scientific Committee on the Effects of Atomic Radiation (UNSCEAR), but national estimates have also been published (Mäkeläinen, Moisio, Reisbacka, & Turtiainen, 2005; United Nations Scientific Committee on the Effects of Atomic Radiation, 2010). The conversion of absorbed dose rate to equivalent RBM dose rate requires data on the subjects' age to model the proportion and location of RBM (Kendall, Fell, & Harrison, 2009).

The RBM dose from indoor radon has not been reported in recent studies as the estimation involves marked uncertainties (Del Risco Kollerud, Blaasaas, & Claussen, 2014; Hauri et al., 2013; Raaschou-Nielsen et al., 2008). Studies have mostly reported estimates based on average indoor radon concentration or its time-integral representing cumulative exposure. Efforts to develop methods to quantify the RBM dose from inhaled radon have been made (Harley & Robbins, 1992; Kendall & Smith, 2005; Richardson, Eatough, & Henshaw, 1991). In fact, radon itself is almost completely exhaled but its daughter nuclei (218-Po and 214-Po), with significantly shorter half-lives, decay in the lungs. Also, the polonium atoms cling to dust particles

due to being more electrically charged and thus get inhaled into the lungs in larger quantities (National Research Council, 1999; World Health Organization, 2009).

1.2.4.3 Average doses in Finland

The average annual radiation dose in Finland has been estimated by STUK – Radiation and Nuclear Safety Authority at regular intervals. In their latest publication, they reported an annual average effective dose of 3.2 mSv (Muikku, Bly, Lahtinen, et al., 2014) (Figure 2). Over half of this dose is from an alpha-emitter, residential radon. Radiation from terrain and rock-based building materials is equal to 0.45 mSv and cosmic radiation in relatively flat Finland has the effective dose of 0.33 mSv. Ingested radioactive elements contribute a dose of 0.32 mSv. The largest man-made part is iatrogenic(0.45 mSv). The slowly disappearing dose resulting from the Chernobyl nuclear accident fallout was approximately 0.02 mSv (Muikku, Bly, Lahtinen, et al., 2014).

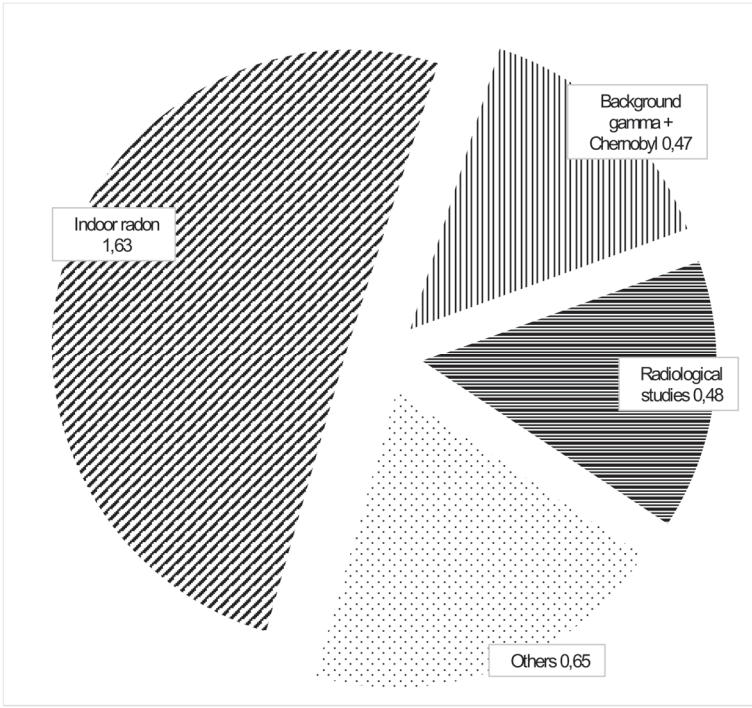


Figure 2. Annual average effective doses in Finland 2012 (Muikku et al., 2014)

The worldwide annual average dose reported by UNSCEAR in 2008 was 3.1 mSv. A slightly lower dose was reported for indoor radon (1.26 mSv) when compared to the average dose from inhaled radon in Finland. (United Nations Scientific Committee on the Effects of Atomic Radiation, 2010) The measured indoor radon concentrations in Finland are very high in comparison to other countries and concentrations above 10 000 Bq/m^3 have also been measured. The average concentration in Finland is 120 Bq/m^3. In Europe, high average concentrations are found also in the Czech-Republic (140 Bq/m^3), Albania (140 Bq/m^3), and Estonia (120 Bq/m^3), as especially low values are found in the Netherlands (23 Bq/m^3), United Kingdom (20 Bq/m^3) and Cyprus (7 Bq/m^3) (United Nations Scientific Committee on the Effects of Atomic Radiation, 2000). The worldwide annual average dose from radiological studies was almost 40% higher (0.62 mSv) than the Finnish average. The number of CT scanners in Finland per million population is relatively low (15.2) in comparison to other developed countries such as Japan (92.6) (United Nations Scientific Committee on the Effects of Atomic Radiation, 2010). The indications, especially for pediatric radiological imaging studies, have been strict in Finland and a shift towards magnetic resonance imaging, which does not utilize ionizing radiation, started around the year 2000 in Finland (Suutari, 2015). An ongoing debate on the number of unnecessary radiological studies on children has probably contributed to increase in the use of magnetic resonance imaging (MRI) (Oren, Kebebew, & Ioannidis, 2019).

2 REVIEW OF THE LITERATURE

2.1 Etiology of childhood leukemia

The pioneering studies on the etiology of childhood leukemia include the twin studies and the studies neonatal blood spots that, for the first time, shed light on the initiating genetic events of the disease (M. Greaves, 2018; Inaba et al., 2013). Also, a plethora of different environmental risk factors from electromagnetic fields to folate deficiency have been suggested and explored to date but still only few have been confirmed (Eden, 2010). Strong evidence exists for high-dose ionizing radiation, certain chemotherapy agents and benzene (Hsu et al., 2013; Pui & Relling, 2000; Carlos-Wallace, Zhang, Smith, Rader, & Steinmaus, 2016). Furthermore, abnormal immunological response to common infectious agents appears to be of importance for transforming pre-leukemic lymphocytes into overt disease (Greaves, 2006; Kinlen, 2012).

2.1.1 Ionizing radiation

High doses of ionizing radiation have been systematically and strongly linked to the genesis of childhood leukemia. The meticulously reported studies on Japanese atomic bomb survivors, for the full cohort, show clear trends for both dose and age at exposure, higher doses and younger children being associated with higher leukemia risk (Hsu et al., 2013). For the cases of ages between 0 and 19 years for a dose of 1 mGy, Hsu et al. (2013) reported a RR of 7.5 (95% CI 5.0, 11.3) for any leukemia excluding chronic lymphoblastic leukemias and adult T-cell leukemias. Based on their analyses a slightly concave dose-response function appeared to fit the risk better than a linear no-threshold model (LNT) (Hsu et al., 2013). A recent review by Boulton covered the subject accurately and it concludes that "exposure to ionizing radiation is an established hazard but it is difficult to gauge the precise risk at less than 100 mSv" (Boulton, 2019). In a pooled analysis of seven historical cohort studies, a significant dose-response was found for doses below 50 mGy for ALL in subjects under the age of 21 years (Little et al., 2018).

Using a case-control design, Stevens et al. studied the nuclear fallout in the Utah area and they report highest excess risks for subjects under 20 years of age at exposure. Using the subset of subjects with doses under 2.9 mGy as the reference group they report for acute leukemias an OR of 1.72 (95% CI 0.94, 3.12) for the category with the highest exposure (>6 mGy) and in subgroup analyses they observed higher estimates related to the youngest ages (Stevens et al., 1990). The Chernobyl fallout has been thoroughly studied but no consistent increase in childhood leukemia has been reported in Finland and the effect was shown to be limited to a maximum of eight excess cases based on the upper 95% confidence interval (Auvinen et al., 1994). In an updated publication. Auvinen et al. state that they lacked sufficient statistical power to analyze children separately (Auvinen et al., 2014). In a Ukrainian study, an increased risk (OR = 4.4, 95% CI 1.3, 15.1 for a comparison of <2.9 mGy and >100 mGy doses) for 0–5-year-old children was observed (Noshchenko, Bondar, & Drozdova, 2010). However, the results have been suggested to be driven partly by sampling bias as similar trends were not seen in Belarus or Russia. A UNSCEAR report also concludes that as no evident risk was present in directly neighboring areas, studies of leukemia at greater distances are unlikely to provide additional useful information (United Nations. Scientific Committee on the Effects of Atomic Radiation, 2013). A recent study on the Ukrainian cleanup workers (adult population) reported elevated SIRs for the combined category of lymphomas and leukemias (Bazyka et al., 2018).

The vicinity of nuclear plants has been studied extensively regarding the incidence of leukemia. However, the conclusion appears to be that there is no clear association between radiation dose and childhood leukemia (Jablon, Hrubec, & Boice, 1991). A recent KiKK study from Germany reported a significantly increased risk related to closer distance to nuclear reactors when comparing distances under
5 km to those living farther away (OR = 2.19, lower one-sided 95% confidence limit 1.51). The subject has been extensively reviewed earlier by Laurier et al. and they conclude that in the vicinity of three nuclear installations (Sellafield, Dounreay and Krümmel) exist an excess of childhood leukemia but the risk could not be validated in multisite studies (Laurier, Grosche, & Hall, 2002). The positive findings have been suggested to be due to infectious etiology according to the delayed infection hypothesis. A more recent detailed review and meta-analysis by Mueller et al. found no increased childhood leukemia risk associated with a distance under 25 km from a nuclear power plant for subjects under the age of 15 years (n=1655). But when they looked at a subgroup of children under the age of five years living closer than 5 km away from a nuclear power plant only from cohort studies, they observed statistically

significant risks (Mueller & Gilham, 2015). In general, the publications on nuclear power plant vicinity and childhood leukemia lack dose estimates and thus they are not highly informative.

The incidence of leukemia around uranium milling and mining sites has been studied and Boice et al. reported a non-significantly increased risk of leukemia (all ages) in Texas (RR = 1.15, 95% CI 0.90, 1.60) (Boice, Mumma, Schweitzer, & Blot, 2003). During the same year they published another article on a similarly conducted study in Pennsylvania reporting a similar increase in leukemia incidence (SIR = 1.45, 95% CI 0.86, 2.30) (Boice, Bigbee, Mumma, & Blot, 2003). A study from New Mexico with a similar setting reported estimates also for a subgroup of subjects below the age of 20 years and a SIR of 1.63 (95% CI 0.91, 2.69) was observed (Boice, Mumma, & Blot, 2010). The role of industrial tritium releases has been studied in an ecological setting and no consistent increase of childhood leukemia was observed related to higher tritium releases when two plants from Germany and the USA were studied (Grosche et al., 1999). These studies were limited by lack of data on dose estimates.

Earlier, large skin hemangiomas of small children were treated with radiation therapy and cohorts have been formed from treated patients and some results suggestive of higher leukemia risks have been reported. Lundell et al. studied a cohort of 14 624 children who were given radiation therapy for hemangiomas in infancy, but they did not observe any statistically significant excess of childhood leukemia (Lundell & Holm, 1996). Also, tinea capitis used to be treated with ionizing radiation. Ron et al. reported a RR of 5.3 (95% CI 1.0, 15) for 1 Gy of dose following the radiotherapy. The authors assumed an average dose of 0.3 Gy from radiotherapy (Ron, Modan, & Boice, 1988). Similarly, a cohort of 2224 children treated for tinea capitis with radiation therapy between 1940 and 1959 were followed and a SIR of 3.2 (95% CI 1.5, 6.1) was reported (Shore, Moseson, Harley, & Pasternack, 2003).

The role of radiation therapy and the risk of developing secondary leukemia has been studied extensively. Tucker et al. had a cohort of 9170 survivors of childhood malignancies, and they observed no increased relative risks from radiation therapy (Tucker et al., 1987). Rosenberg et al. made similar observations few years earlier after a 22-year follow-up of Hodgkin's disease patients (Rosenberg & Kaplan, 1985) and they concluded that irradiation does not appear to be major factor regarding secondary AMLs after Hodgkin's disease therapy. However, Hawkins et al. reported results from a case-control study with 26 survivors of childhood cancer with secondary leukemia as cases and they observed a significant dose-response trend for the radiation therapy dose and risk of secondary leukemia ($p=0.012$) (Hawkins et al.,

1992). Positive results were reported also by Kaldor et al. as they used a case-control design in international collaboration with 163 cases of Hodgkin's disease survivors. They reported a RR of 1.24 (95% CI 1.04, 1.43) for every 1 Sv of dose to the red bone marrow (Kaldor et al., 1990). More recently, Friedman et al. reported results after a follow-up of 14 359 North-American childhood cancer survivors. They observed an excess of leukemia diagnoses (SIR = 6.1, 95% CI 4.5, 8.2) and in a multivariate analysis they observed an independent elevated risk of any secondary neoplasm in the presence of any radiation therapy (RR = 2.7, 95% CI 2.2, 3.3). Leukemia diagnoses were not analyzed separately (Friedman et al., 2010). Haddy et al. observed, however, an elevated non-significant risk (Haddy et al., 2006). They reported a RR of 4.2 (95% CI 0.8, 20.7) for secondary leukemia after radiotherapy for solid tumors. The RBM doses from the radiation therapy ranged from 3.0 Gy to 6.6 Gy.

The childhood leukemia risk has been assumed to follow a linear no-threshold model, which is created by extrapolating from risk estimates calculated for higher doses of ionizing radiation (Shore et al., 2018). It has been emphasized that the use of a linear model is essentially a default approach due to lack of definitive evidence, but the available epidemiologic publications support its use and before new results, it remains a pragmatic choice (Shore et al., 2018). Nevertheless, its use for smaller (<100 mGy) doses has been criticised and other shapes have also been suggested for the low end of the dose-response curve (Aurengo, Averbeck, & Bonnin, 2005; Hsu et al., 2013; United Nations Scientific Committee on the Effects of Atomic Radiation, 2010). A slightly concave dose-response curve, which is achieved by adding a quadratic term to the function, fits the LSS data better (Hsu et al., 2013). Even hormetic effects, below a certain threshold, have been suggested (Shibamoto & Nakamura, 2018). However, these claims have gained little to no support.

Regarding specifically adult leukemia and ionizing radiation, radiation monitored workers have been studied extensively and a large international cohort (INWORKS) reports an excess risk of leukemia (RR = 2.96, 90% CI 1.17, 5.21 per 1 Gy of red bone marrow dose). They used a two-year lag period in their analyses, however, they state that due to classical measurement error, the point estimates could be diluted. Their results linearly translate to a RR of 1.003 for every 1 mGy of red bone marrow dose, which is of expected magnitude for the adult population (Leuraud et al., 2015). Also, Techa river area residents were exposed to elevated doses of bone marrow seeking Sr^{90} radionuclide. The doses reached levels exceeding 1 Gy. Leukemias other than CLL showed a RR of 1.22 (95% CI 1.08, 1.54) for every 100 mGy dose increase

(Krestinina et al., 2013). The heterogeneity between different age groups was also tested but no statistically significant difference was observed.

2.1.1.1 Background gamma radiation

Background radiation can be defined to consist of components originating from terrestrial uranium, cosmic radiation and remnants of nuclear disasters and tests. Its effects on childhood leukemia have been studied but the results of studies have been conflicting (Table 2).

Table 2. Studies on background radiation and childhood leukemia

Authors	Country	Publication year	Study design	Time period	Follow-up duration (y)	Mean exposure	Latency period (y)	Confounders adjusted for	N (cases)	Age (y)	Risk
Spix et al.	Germany	2017	Ecologic	1987–2011	*not applicable*	*not applicable*	*not applicable*	Age	13,374	<15	RR = 1.04 (95% CI 0.91, 1.20), *0.5 mSv/year vs. 1.5 mSv/year for ALL*
Demoury et al.	France	2017	Case-control	1990–2009	*not applicable*	98.2 nSv/h	0, 2	Age, nuclear plants, high-voltage power lines, proximity to high traffic roads	9056	<15	SIR = 1.00 (95% CI 0.99, 1.01) *for 1 mSv increase in cumulative RBM dose*
Spycher et al.	Switzerland	2015	Cohort	Censuses at 1990 and 2000	7.7	109 nSv/h	0	Sex, birth year, proximity to high traffic roads, electromagnetic fields, high-voltage power lines, urbanization, socioeconomic status, parental education, number of persons per room, birth weight, birth order	530	<16	HR = 1.04 (95% CI 1.00, 1.08) *for 1 mSv increase in cumulative effective dose*
Kendall et al.	United Kingdom	2013	Case-control	1980–2006	*not applicable*	94.7 nSv/h	0, 0.75, 1, 2	Age, sex, socioeconomic status	9058	<15	RR = 1.12 (95% CI 1.03, 1.22) *for 1 mSv increase in cumulative RBM dose*
UKCCS	United Kingdom	2002	Case-control	1991–1996	*not applicable*	96.1 nGy/h	0	Age, sex, socioeconomic deprivation	2165	<15	OR = 0.95 (95% CI 0.66, 1.37) *for the highest quintile of average dose rate (≥1045.3 μSv/year)*

First, a United Kingdom Childhood Cancer Study did not report increased risk related to higher annual doses of background radiation (UK Childhood Cancer Study Investigators, 2002b). They had access to direct measurements of gamma and cosmic radiation and included only patients who had lived a minimum period of six months in their dwellings. The analyses were adjusted for deprivation index as the authors had observed a correlation between higher exposures and higher socioeconomic deprivation index.

Almost ten years after, a large study, also from the United Kingdom, found a significantly increased risk related to natural background radiation (Kendall et al., 2013). The study by Kendall et al. (2013) was able to avoid many shortcomings of the earlier study. In particular, the study had significantly more statistical power, which is necessary when studying exposures of small expected effect size. Unfortunately, the case-control data was matched to controls by location due to use of local birth registries, and in the regression analyses of dose rate, they lost 48% of cases as they had identical dose rates to their controls.

A Swiss cohort study replicated the positive finding of Kendall et al. with a slightly lower risk estimate by using a cohort design (Spycher et al., 2015). They identified children under the age of sixteen from two national censuses (1990 and 2000) and their follow-up ended in 2008. Their dose estimates included cosmic radiation, natural terrestrial radiation and artificial terrestrial radiation in a 2 x 2 km grid map. The natural terrestrial radiation modelling was based on measurements carried out with a helicopter, in situ gamma-ray spectrometry, in situ dose rate measurements with ionization chambers and laboratory measurements. Furthermore, this study was able to adjust for a large number of confounding factors.

In contrast, this was followed by a large French study with no signs of an association (Demoury et al., 2017). The French study used indoor gamma measurements from dental clinics as the basis for dose estimation and they were able to include cosmic radiation in their analyses. They constructed a 1 x 1 km grid and estimated that their exposure assessment approach explained 65% of the variance by spatial coordinates.

In line with the Demoury et al. study, a large ecological German study did not report statistically significant results related to higher doses, however, the central estimate of their analyses was above unity and of the expected effect size (Spix et al., 2017). They used a fractional polynomial model as the basis of their analysis and their radiation data was based on extrapolation from 1800 outdoor measuring stations. As part of exploratory subgroup analyses, they reported a non-significant

association related to the age group of subjects under one year old (RR = 1.42, 95% CI 0.91, 2.21).

In general, the results from all listed studies are in agreement and the confidence intervals include a risk of ~1.01/mSv linearly approximated from the estimates of life-span studies (Hsu et al., 2013). Interestingly, Wakeford et al. estimated that approximately 20% of childhood leukemia cases in the United Kingdom could be explained by background radiation (Wakeford, Kendall, & Little, 2009).

In discordance, studies on populations living in areas with naturally very high background radiation have not reported elevated risks. A study was conducted with an Indian cohort living in the Kerala area where, at certain locations, background radiation dose rates reach almost 10 000 nGy/h and the median dose rate is 456 nGy/h (Nair et al., 2009). Even with such high dose rates, they did not observe a significant excess risk related to leukemia (RR = 1.42, 95% 0.36, 5.55 when comparing cumulative dose of <50 mGy with doses of >200mGy). But this study was conducted with subjects aged over thirty years. Another study on adults above 30 years of age from China did not report positive findings related to leukemia when using a two-year latency period (Tao et al., 2012). The adult studies need to be interpreted carefully in the context of childhood leukemia as the etiologies are likely to be different.

2.1.1.2 Diagnostic radiological examinations

The effect of radiological studies on childhood leukemia has been explored extensively prior to the studies on computerized tomography. An early study from New Zealand reported an increased risk of leukemia after therapeutic irradiation for RBM doses over 100 mGy, but no clear association was found for diagnostic radiation (Gunz & Atkinson, 1964). They did not report results by age groups. Graham et al. reported an increased non-significant risk (RR = 1.14, p>0.5) for post-natal diagnostic imaging studies using a lag of one year (Graham et al., 1966; United Nations Scientific Committee on the Effects of Atomic Radiation, 2013). Childhood leukemia has also been studied separately in a similar setting by Bartley et al. using a lag of one year and they found a significant association for post-natal diagnostic imaging studies for ALL for three or more diagnostic X-rays (OR = 1.85, 95% CI 1.12, 2.79) but no significant association was reported for AML (Bartley, Metayer, Selvin, Ducore, & Buffler, 2010). Infante-Rivard reported similar findings to Bartley et al. and they reported an OR of 1.48 (95% CI 1.11, 1.97) for two or more x-rays (Infante-Rivard, 2003).

However, there are also multiple studies with null findings regarding plain X-ray imaging. Meinert et al. did not observe any elevated risks for diagnostic X-ray examinations for children under the the age of 1 year using a case-control design (Meinert, Kaletsch, Kaatsch, Schüz, & Michaelis, 1999). However, they reported results also for paternal imaging studies two years prior to birth and for an imaging study of any site they observed a significant OR of 1.33 (95% CI 1.10, 1.61). For maternal examinations before or during pregnancy, no elevated risks were reported. Another study reported an OR of 1.36 (95% CI 0.91, 2.02) for childhood leukemia (Rajaraman et al., 2011) for *in utero* exposure to X-ray examinations. In their study, the risk was more pronounced for AML (OR = 2.44, 95% CI 0.95, 6.33). Shu et al. studied parental X-ray exposure preconception, during pregnancy (for females only) and for children also post-natally (Shu et al., 2002). Overall, they did not report any statistically significant findings, but they observed an elevated risk for precursor B-cell acute lymphoblastic leukemia in relation to post-natal diagnostic X-ray examinations.

Computerized tomography scans yield markedly higher radiation doses than plain film radiography and its increasing use in developed countries, especially in the pediatric population, has raised concerns. Higher radiation doses are associated with higher risks and, in some cases, the indications of CT scans could be reconsidered. As a guiding principle, the increased risks should always outweigh the clinical benefits (Brenner, 2007). A typical brain CT scan of a 10-year-old child after 2001 has been estimated to yield a RBM dose of 6.0 mGy whereas a typical skull x-ray yields an effective dose of approximately 0.03 mGy (Kim et al., 2012; Wall & Hart, 1997). The use of computerized tomography scans increased after its introduction in the early 1980's but, in Finland, due to the higher radiation doses, the pediatric use of CT scanning started to decline as early as after the year 2000 as MRI became more common (Suutari, 2015). In many cases, CT scans provide valuable diagnostic information, but valid concerns are founded regarding its increased use due to the delivered doses of ionizing radiation (Oren et al., 2019).

The effects of computerized tomography scans on childhood leukemia have been studied mostly using large cohort studies but concerns about the effects of predisposing factors have risen as they might result in confounding by indication (Table 3). Predisposing factors are conditions associated with probabilities of both computerized tomography scans and childhood leukemia, e.g. Down syndrome (DS).

Table 3. Cohort studies on computerized tomography scans and childhood leukemia

Authors	Country	Year	Time period	Follow-up duration (y)	Mean dose (mGy)	Latency period (y)	Confounders adjusted	Exposed cases	Age (y)	Risk
Meulepas et al.	The Netherlands	2018	1979–2012	8.5	9.5	2	Household income	44	<18	RR = 1.00 (95% CI 0.99, 1.03) for 1 mGy RBM dose
Berrington de Gonzalez et al.	United Kingdom	2016	1985–2002	9.6	18.9 for cases	2	Sex, Socioeconomic status, time since first exposure	74	<22	RR = 1.03 (95% CI 1.00, 1.12) for 1 mGy of RBM dose when leukemia related conditions were excluded, RR = 1.02 (95% CI 0.99, 1.09) for 1 mGy of RBM dose when subjects with previous cancers were excluded
Krille et al.	Germany	2015	1980–2010	4.1	11.7	2	Sex, age	12	<15	HR = 1.01 (95% CI 0.98, 1.04) for 1 mGy of RBM dose
Journy et al.	France	2015	2000–2010	4.4	6.9 (median)	1, 2, 3, 4	Sex, period of birth, age, time in cohort	25	<10	RR = 1.05 (95% CI 0.93, 1.16) for 1 mGy of RBM dose
Huang et al.	Taiwan	2014	1998–2006	4.4	Not reported	2	Sex, age, other than head CT	25	<18	HR = 1.90 (95% CI 0.82, 4.40), 1 or more scans vs. no CT scans
Mathews et al.	Australia	2013	1985–2005	9.5	4.6	1, 5, 10	Age, sex and year of birth	211	<20	RR = 1.04 (95% CI 1.00, 1.08) for 1 mGy of mGy of RBM dose
Pearce et al.	United Kingdom	2012	1985–2002	9.6	18.9 for cases	2	Sex, Socioeconomic status, time since first exposure	74	<22	RR = 1.04 (95% CI 1.01, 1.12) for 1 mGy of RBM dose

A large cohort study from the United Kingdom reported an increased dose-response gradient with RBM dose (Pearce et al., 2012). However, they included myelodysplastic syndromes in their analyses, which mostly accounted for the elevated risk estimates. Their dose estimates were based on two nationwide surveys (1989 and 2003) and data from a series of hybrid computational human phantoms. The RBM doses before the year 2001 were on average over two times higher.

One year later, the British study was followed by a large Australian cohort study also reporting a similar risk estimate per 1 mGy (Mathews et al., 2013). As the preceding study, the authors could not obtain individual scanning parameters and their dose estimates were based on scanning site, year of the scan and the subject's age. Also, in their analyses, the doses dropped markedly after year 2001 as scanning parameters started to be adjusted based on a patient's size and age to reduce radiation doses. The authors recognize themselves that they could not attribute the observed excess of leukemias completely to computerized tomography scans as they could not adjust for predisposing factors.

A much smaller Korean study reported a non-significantly increased risk for 1 or more CT scans (Huang et al., 2014). The study focused solely on head CT scans and they did not attempt to model red bone marrow radiation doses at all. They did not have data on the indications of the head CT scans and, thus, they recognize a possibility of screening bias. However, they had diagnostic data on precancerous conditions and by exclusions they tried to minimize the effect of predisposing syndromes.

The issue of unknown predisposing factors may have biased many of the studies on childhood leukemia and CT scans (Journy et al., 2015). Especially the Australian study was criticized for this reason (United Nations. Scientific Committee on the Effects of Atomic Radiation, 2013). The presence of predisposing factors could result in overestimations of the actual childhood leukemia risk related to pediatric CT scans. The French study utilized a software-based age-specific approach (MCNPX 2.7) in their dose estimates and for 13% of the scans they had to use median doses due to missing data. The authors approached the predisposing factors with a list of diagnoses constructed based on a literature review by local experts in pediatric oncology and they searched for the diagnoses from the discharge databases of the included hospitals.

A German study did not observe a significantly increased risk related to one CT scan (SIR = 1.18, 95% CI 0.43, 2.57) but the results for two or more CT scans were statistically significant (SIR = 3.17, 95% CI 1.16, 6.90) (Krille et al., 2015). They did not observe a statistically significant risk per 1 mGy (HR = 1.01, 95% CI 0.98, 1.04).

They based their dose estimates on a previously published catalogue as no individual dosimetry was available (Miglioretti et al., 2013). The authors manually reviewed radiology reports and, in total, excluded 15 CT scanned children as hints of reverse causation and confounding by indication were observed.

Berrington de Gonzalez et al. reanalyzed the data previously reported by Pearce et al. and after accounting for multiple potential sources of bias in several scenarios their results remained largely similar (Berrington De Gonzalez et al., 2016). The cancer-predisposing syndromes were collected from three separate sources. For approximately a third of the patients they reviewed hand-written comments in the radiology reports and for all the deceased subjects, the causes of death were reviewed. Also, pathology reports related to leukemia or myelodysplastic syndrome diagnoses were reviewed.

A recent Dutch study did not find significantly elevated risks (Meulepas et al., 2018). Contrary to previous studies, they were able to collect accurate information on radiation doses for approximately 40% of the included CT scans and the remaining 60% were estimated using the NCICT software. They were able to approximate that their methods explained 80% of the variation in the RBM doses. They were not able to collect data on clinical indications of the CT scans and only tuberous sclerosis diagnoses were retrieved in order to address potential predisposing factors.

Overall, the results favor a risk associated with childhood leukemia. However, in many studies the risks were higher than expected (RR = 1.01/mSv) based on the Japanese LSS studies and still the relationship between imaging indications and childhood leukemia should be emphasized to avoid confounding by indication (Hsu et al., 2013).

2.1.1.3 Indoor radon gas

Indoor radon concentrations vary markedly and differences of multiple orders of magnitude are not uncommon (Mäkeläinen et al., 2009). The construction tradition in Finland differs from most of the inhabited world due to the cold climate. Thus, national studies to investigate important determinants of higher indoor radon concentrations are needed. The indoor radon concentration depends on a number of attributes of the building and the soil beneath it (Mäkeläinen et al., 2009). Also, lifestyle differences between individuals living in a dwelling play a role.

An important determinant of indoor radon concentration is the type of the dwelling as they are built quite differently. Commonly, low rise residential buildings

such as terraced houses, semi-detached houses and single-family houses are called houses. Dwellings in blocks of flats are referred to as apartments. Houses differentiate themselves clearly from apartments in Finland by the attributes determining indoor radon concentrations.

In houses, the most important source of radon is the soil-borne gas as the buildings are in contact with the ground. The uranium concentration of the underlying soil determines the rate of radon production but, after that, the permeability of the soil allows gases to rise and enter the dwelling. The foundation of the building acts as the barrier between the soil and the living spaces of the dwelling and, thus, all its structures allowing the movement of radon gas (gaps, gas permeable blocks, openings in structures) are associated with higher indoor radon concentrations (Mäkeläinen, Arvela, & Voutilainen, 2001) (Figure 3). However, a study by Revzan et al. showed that the permeability of the soil just beneath the building is a more important determinant of the indoor concentrations than the combined area of the leakage routes (Revzan & Fisk, 1992). The highest concentrations are, therefore, measured in dwellings located on coarse gravel and, conversely, the ones situated on impermeable clay have the lowest concentrations. Also, type of ventilation is an important determinant (Arvela, Holmgren, Reisbacka, & Vinha, 2014). In addition, a lower air pressure indoors allows higher flow rates of radon gases through the entry routes. Thus, well-insulated and relatively air tight warm Finnish houses equipped with only exhaust fans create optimal conditions for higher indoor radon concentrations to start to accumulate. It has also been shown that the existence of cellars is associated with higher concentrations (Barros-Dios, Ruano-Ravina, Gastelu-Iturri, & Figueiras, 2007; Mäkeläinen, Valmari, Reisbacka, Kinnunen, & Arvela, 2010). There are also seasonal variation in indoor radon concentrations (Mäkeläinen et al., 2009). During summer, the houses are typically more ventilated, and this dilutes the indoor radon concentration. The practical guidelines for construction of new buildings were updated in 2004 and in dwellings built thereafter, measures to decrease indoor radon concentrations have been taken. A Finnish survey shows that houses with radon protection, have approximately 50% lower concentrations on average (Arvela, Holmgren, & Reisbacka, 2012).

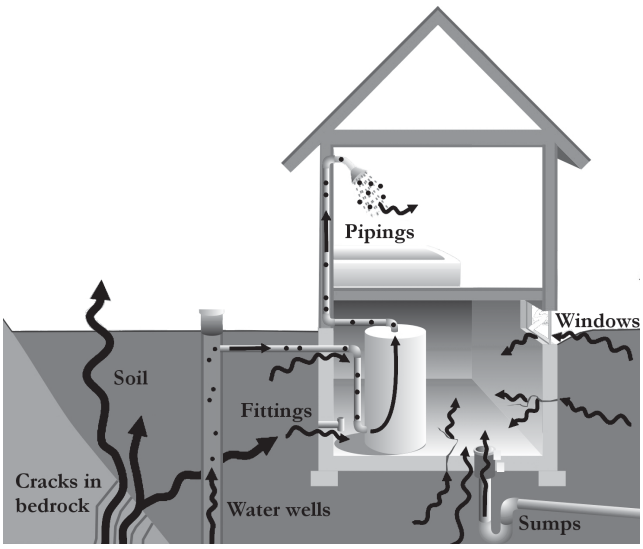


Figure 3. Illustration of different possible entry routes for radon gas. Adapted from an illustration by Natural Resources Canada 2008, Geological Survey of Canada.

In apartments, only the ground floor is in contact with the soil and thus soil-borne radon is not as important. The storey of the dwelling has been reported to be a significant predictor of higher concentrations (Lorenzo-González, Ruano-Ravina, Peón, Piñeiro, & Barros-Dios, 2017, p.; Valmari, Arvela, & Reisbacka, 2012). In other floors, the radon is formed in the rock-based building materials. When modelling the radon emanating from building materials, it is reasonable to assume that the local gravel material has been used in production of the concrete elements as is typical in Finland. In general, the average indoor radon concentration is lower in apartments (49 Bq/m^3) when compared with the average concentration in houses (121 Bq/m^3) (Mäkeläinen et al., 2009).

In recent studies of childhood leukemia and indoor radon from Denmark and Switzerland without area-based prediction models, important final predictors included dwelling type, floor, basement, soil geology, region, building material, soil permeability, degree of urbanization and the year of construction (Andersen et al., 2007; Hauri, Huss, Zimmermann, Kuehni, & Röösli, 2012). The Danish model was based on only 3116 measurements (91 apartments and 3025 houses) and the Swiss model was based on 35 706 measurements (the numbers of different housing types not reported).

The mechanism, by which inhaled radon gas could affect the red bone marrow, has been debated. Harley and Robbins estimated that in Denmark an additional 10 to 60% of red bone marrow dose could be attributed to inhaled radon (Harley & Robbins, 2009). However, they also investigated the doses deposited to pulmonary lymphocytes residing among basal and mucous cells of the alveoli and bronchioles and ended up with an unexpectedly large approximation of 100 mSv annual dose from continuous exposure to a radon concentration of 100 Bq/m^3 (Harley & Robbins, 2009).

High concentrations of inhaled radon gas have been linked with increased risk of lung cancer in recent meta-analyses (Darby et al., 2005; Garzillo, Pugliese, Loffredo, & Quarto, 2017; Krewski et al., 2006). A Czech uranium miner study suggestive of leukemia risk in adults was published by Reiricha et al. by observing 23 043 miners with 177 incident cases (Rericha, Kulich, Rericha, Shore, & Sandler, 2006). Similar results have been published from a German miner cohort as an increased risk for non-CLL was observed in relation to higher radon concentrations (Kreuzer, Sobotzki, Fenske, Marsh, & Schnelzer, 2017). On adult leukemia, Law et al. published a short report with null findings on the association between residential radon and leukemia (Law, Kane, Roman, Smith, & Cartwright, 2000). They reported quintiles of measured radon, and higher categories did not differ from unity more than the lower ones. The effect of indoor radon on childhood leukemia has been studied extensively but the results have been conflicting (Table 4).

Table 4. Studies on association of indoor radon and childhood leukemia with individual level data

Authors	Country	Year	Study design	Follow-up duration (y)	Mean exposure	Latency period (y)	Confounders adjusted	Exposure assessment	Time period	N (cases)	Age (y)	Risk
Demoury et al.	France	2017	Case-control	*not applicable*	67.8 Bq/m^3	0, 2	Age, nuclear plants, high-voltage power lines, proximity to high traffic roads	Area-based	1990–2009	9056	<15	SIR = 1.01 (95% CI 0.91, 1.12) for every 100 Bq/m^3 increase in average concentration
Kollerud et al.	Norway	2014	Cohort	12.4	91 Bq/m^3	0	parental education and family income, sex, birth weight, congenital malformations, parity	Model-based	1967–2009	431	<16	HR = 1.00 (95% CI 0.87, 1.15) for every 100 Bq/m^3 increase in average concentration
Hauri et al.	Switzerland	2013	Cohort	5.7	86 Bq/m^3	0	Age, sex, birth order, parental socioeconomic status, gamma radiation, time period	Model-based	1984–2008	283	<16	HR = 0.90 (95% CI 0.68, 1.19) for every 100 Bq/m^3 increase in average concentration
Kendall et al.	United-Kingdom	2013	Case-control	*not applicable*	21.3 Bq/m^3	0, 0.75, 1, 2	Age, sex, socioeconomic status	Area-based	1980–2006	9058	<15	RR = 1.12 (95% CI 0.88, 1.43) for every 1000 Bq/m^3-year increase in cumulative exposure
Raaschou-Nielsen et al.	Denmark	2008	Case-control	*not applicable*	48 Bq/m^3	0	Mother's age, birth order, traffic density, high-voltage power lines	Model-based	1968–1994	2400	<15	OR = 1.70 (95% CI 1.08, 2.67) for every 1000 Bq/m^3-years increase in cumulative exposure for ALL
Yoshinaga et al.	Japan	2005	Case-control	*not applicable*	18 Bq/m^3	0	magnetic fields, maternal education, maternal X-ray exposure during pregnancy	Direct measurement	1999–2002	227	<15	OR = 1.57 (95% CI 0.47, 5.22) for <20 Bq/m^3 vs. 50–100 Bq/m^3
UKCCS et al.	United-Kingdom	2002	Case-control	*not applicable*	24 Bq/m^3	0	Age, sex, socioeconomic status, double glazing in windows, central heating, study region	Direct measurement	1991–1996	2226	<15	OR = 0.95 (95% CI 0.49, 1.84) for <24 Bq/m^3 vs. >200 Bq/m^3
Maged et al.	Egypt	2000	Case-control	*not applicable*	55 Bq/m^3	0	-	Direct measurement	1996–1998	50	2-14	OR = 5.42 (95% CI 1.3, 21.1), <40 Bq/m^3 vs. >90 Bq/m^3 for ALL
Kaletsch et al.	Germany	1999	Case-control	*not applicable*	27 Bq/m^3	0	Age, sex, urbanization, socioeconomic status	Direct measurement	1988–1993	219	<15	OR = 1.30 (95% CI 0.32, 5.33), >70 Bq/m^3 vs. <70 Bq/m^3
Steinbuch et al.	USA and Canada	1999	Case-control	*not applicable*	54 Bq/m^3	0	Age, sex, maternal education, family income, maternal ethnicity, maternal age, number of hours spent away from home	Direct measurement	1989–1993	638	<18	OR = 1.1 (95% CI 0.6, 2.0), >100 Bq/m^3 vs. <37 Bq/m^3 for AML
Lubin et al.	USA	1998	Case-control	*not applicable*	79.1 Bq/m^3	0, 2	Age, sex	Direct measurement	1989–1993	281	<15	RR = 1.02 (95% CI 0.5, 2.0), >148 Bq/m^3 vs. <37 Bq/m^3 for ALL

There are numerous ecological studies exploring residential radon gas and childhood leukemia with mainly positive results, but the setting is inherently more prone to bias when compared to observational studies (Henshaw, Eatough, & Richardson, 1990; Richardson et al., 1991; Tong et al., 2012).

Studies on indoor radon and childhood leukemia with direct indoor radon concentration measurements have not reported statistically significant increases in leukemia risk. A North-American study by Lubin et al. started building on earlier ecological studies with mainly positive results (Lubin et al., 1998). They reported no association between indoor radon exposure and ALL (Lubin et al., 1998). The authors excluded patients who had lived in the measured dwelling for less than six months and also required a minimum measurement coverage of 70% for the last five years preceding the reference dates of the study subjects.

A year later a similar study on AML was published and the authors found no association related to AML (Steinbuch, Weinberg, Buckley, Robison, & Sandler, 1999). However, they observed an increased risk related to an age group consisting of children older than two years (OR = 2.79, 95% CI 1.0, 7.8), a finding the authors attributed to chance. The authors were able to model the time spent indoors but their results remained largely similar after this was accounted for. To control for confounding they excluded all patients with Down syndrome (N=19).

Kaletsch et al. studied the subject in Germany and with a predefined cutoff value 70 Bq/m^3, they observed no significantly elevated risk related to leukemia (OR = 1.30, 95% CI 0.32, 5.33). However, the average indoor radon concentrations were quite low in the studied dwellings (27 Bq/m^3), which reduced the potential to make deductions for higher concentrations (Kaletsch et al., 1999). Based on telephone interviews, the full residential histories were collected, and radon measurements were carried out in up to three rooms inside the dwellings. All measurements with values below 10 Bq/m^3 were excluded due to suspected artefacts in the measurements.

A small Egyptian study reported markedly higher odds ratio (OR) of 5.42 (95% CI 1.3, 21.1) related to ALL when comparing average exposures of <40 Bq/m3 with those of >90Bq/m3 (Maged et al., 2000). The magnitude of Egyptian findings is likely affected by random error as the main results differ so drastically from the previous studies and are considerably above the expected effect size. Direct indoor radon measurements were carried out for 10% of the initially contacted ALL cases. They also looked for attributes associated with higher indoor radon concentrations and, in their analyses, only ventilation showed a significant association while, for example, floor level showed no association.

The latest large measurement-based study was reported by the United Kingdom Childhood Cancer Study (UKCCS) and they did not observe any elevated risk of childhood leukemia related to higher indoor radon concentrations (OR = 0.95, 95% CI 0.49, 1.84, for the concentrations >200 Bq/m^3) (UK Childhood Cancer Study Investigators, 2002a). For a majority of cases they carried out radon measurements in two different rooms in the dwelling preceding diagnosis by six months. However, their study also suffered from low average radon concentrations (21–25 Bq/m^3).

Preliminary results from a Japanese study with only a tenth of the sample size of the UKCCS were published in 2005 as a congress paper (Yoshinaga, Tokonami, Akiba, Nitta, & Kabuto, 2005). They had only ten cases with exposures exceeding 50 Bq/m^3 and for the second exposure category (20 – 49 Bq/m^3) they reported an OR of exactly one (95% CI 0.62, 1.62). The categories with higher exposure showed a non-significant upward trend but they lacked in sample size.

A dilemma of choosing between a larger sample size and more accurate exposure assessment has prevailed in the field. Evolution of studies on childhood leukemia and indoor radon, in general, started with ecological studies, then moved to case-control studies with direct measurements and then towards model-based studies. A Danish effort on the effect of residential radon was published in 2008 (Raaschou-Nielsen et al., 2008). Raaschou-Nielsen et al. reported an elevated relative risk (RR) of 1.56 (95% CI 1.05, 2.30) per 1000 Bq/m^3 years for ALL. When they restricted the analysis to subjects who had lived in only single-family houses throughout their childhood, the excess risk, interestingly, almost tripled to a RR of 2.44 (95% CI 1.24, 4.81). Importantly, the Danish study had complete register-based residential histories available, which made the exposure assessment more accurate. In addition, the average indoor radon concentrations are low in Denmark (59 Bq/m^3) in comparison to other European countries (United Nations Scientific Committee on the Effects of Atomic Radiation, 2000). Before publishing their study on risk assessment, the prediction model was presented and evaluated separately (Andersen et al., 2007).

A similar model-based approach was used by Hauri et al. with Swiss data (Hauri et al., 2013). Using a census-based cohort study design and a pre-validated prediction model they did not find evidence of an increased childhood leukemia risk (Hauri et al., 2012). In age-adjusted analyses they reported a HR of 0.97 (95% CI 0.74, 1.27) for an increase of 100 Bq/m^3 in average radon concentration.

Kendall et al. also evaluated the effect on indoor radon using area-based radon concentration estimation (Kendall et al., 2013; Miles et al., 2007). In the UK study, a non-significant relative risk of 1.12 (95% CI 0.88, 1.43) related to childhood

leukemia was reported. The geometric mean of indoor radon concentrations in the United Kingdom was low (16.4 Bq/m^3).

A recent publication on childhood leukemia based on Norwegian data reported no elevated risk related to childhood leukemia and in a model adjusted for all available potential confounders observing a HR of 1.00 (95% CI 0.87, 1.15) for every 100 Bq/m^3 increase in average concentration (Del Risco Kollerud et al., 2014). They had direct measurements available only for 6% of the dwellings and they predicted the concentrations for the missing measurements using at least five nearest available measurements inside a defined buffer radius for prediction (Kollerud et al., 2014).

The most recent attempt to study the matter was by a French study and it was published alongside their study on terrestrial background radiation. For all childhood leukemias they reported a SIR of 1.01 (95% CI 0.91, 1.12) for every 100 Bq/m^3 increase in average indoor radon concentration. The exposure assessment was based on 10,843 measurements of indoor radon and a 1 x 1 km grid map was constructed. The authors estimate that their approach was able to explain 32% of the variance in indoor radon based on spatial coordinates of the home location (Demoury et al., 2017).

2.1.2 Other well-established risk factors

Down syndrome has been identified as a risk factor for childhood leukemia with a considerably high-risk estimate. Hasle et al. report a standardized incidence ratio of 17.6 (95% CI 12.4, 24.4) for all leukemias. For children under the age of ten, the SIR was 56.4 (95% CI 37.8, 81.0) (Hasle, 2001).

Rare genetic conditions affecting DNA repair have been associated with childhood leukemia. Alter et al. combined results from four cohort studies and found a relative risk of 700 for AML among Fanconi anemia patients (Alter, 2014). Incidences are severely elevated (>10% of subjects with the predisposition) also for familial platelet disorder, Schwachman-Diamond syndrome, Severe congenital neutropenia (including Kostmann syndrome), Nijmen breakage syndrome, dyskeratosis congenital, Wiskott-Aldrich syndrome and Bloom syndrome among a few others rare genetic diseases (Stieglitz & Loh, 2013).

A meta-analysis by Caughey et al. reported an elevated OR of childhood leukemia for large birth weight (1.35, 95% CI 1.24, 1.48). The OR per 1000g of additional weight was 1.18 (95% CI 1.13, 1.23) and specifically for AML they observed

suggestions of a U-shaped risk association implying a potential risk related also to low birth weights (Caughey & Michels, 2009).

In a review article by Pui et al. literature on chemotherapy agents was reviewed and topoisomerase II inhibitors, some anthracyclines, mitoxantrone, dactinomycin and dioxypiperazine derivatives were reported to be capable of triggering therapy related AML as a secondary malignancy (Pui & Relling, 2000). The evidence was conflicting regarding the existence of dose-response relationship. In summary, cumulative etoposide doses below 2000 mg/m^2 had negligible risk (<1%) of secondary AML (Pui & Relling, 2000).

A review by Belson et al. summarizes the effects of parental occupational exposures and positive results were reported on parental exposure to paints and pigments (Belson, Kingsley, & Holmes, 2007). Leukemia was also reported more frequently for children whose mothers had worked in metal manufacturing, textile industry or drug manufacturing. Positive results were reported also for solvents and pesticides. A study by Buckley et al. included results also on benzene and they reported an OR of 2.4 (95% CI 1.3, 4.1) for over 1000 days of occupational exposure to petroleum products (Buckley et al., 1989). A meta-analysis on parental benzene exposure also reported a relative risk of 1.96 (95% CI 1.53, 2.52) for any parents' occupational or household exposure to benzene (Carlos-Wallace, Zhang, Smith, Rader, & Steinmaus, 2016). The risk estimates were higher for maternal exposure and AML.

Differences in risk of childhood leukemia between ethnicities have been reported in the USA using a large case-control study (5788 cases, 5788 controls) especially related to ALL as the incidence of AML is more similar for all major ethnicities residing in the USA (Oksuzyan et al., 2015). In the USA, the odds ratio for African American children was significantly lower (OR = 0.54, 95% CI 0.45, 0.66) when compared to white children. Also, Hispanic ethnicity was associated with higher risk of ALL (OR = 1.37, 95% CI 1.22, 1.52) as Asian ethnicity was associated with an increased risk of AML (OR = 1.64, 95% CI 1.10, 2.46).

Familial aggregation of childhood leukemia has been studied and a higher than expected percentage of cases was observed for monozygotic twins when compared with dizygotic twins (Buckley et al., 1996). Unlike for other diseases, a different mechanism, instead of shared inheritance, has been suggested for childhood leukemia and the mutation could be disseminated to the other twin via placental circulation (Ford et al., 1993). In general, childhood ALL does not appear to be affected by overall history of cancer (Zierhut, Linet, Robison, Severson, & Spector, 2012). Sibship concordance of childhood leukemia has been studied and specific

subtypes co-occurred significantly more than expected ($p<0.0001$) (Schmiegelow et al., 2011).

2.1.3 Well-established protective factors

Some factors have been found to be protective of childhood cancer. Breastfeeding has shown protective signs in a meta-analysis of 17 studies. Any breastfeeding for a duration longer than 6 months showed an OR of 0.80 (95% CI 0.72, 0.90) and, thus, the authors suggest that breastfeeding infants could, in theory, prevent 14 – 20% of all childhood leukemia diagnoses (Amitay & Keinan-Boker, 2015).

Daycare attendance has also been studied and a significantly reduced effect was observed. The exposure in the study was defined as spent time in daycare with the presence of a given number of children (child-hours). The authors report an OR of 0.991 (0.984, 0.999) related to every 1000 child-hours of daycare (Ma et al., 2002). The findings support the delayed infection hypothesis (Greaves, 2006).

In a pooled analysis of the Childhood Leukemia International Consortium (CLIC) studies, maternal folic acid supplementation at preconception was associated with a lower risk of childhood leukemia (OR = 0.85, 95% CI 0.78, 0.92). In more detailed analyses the authors observed that the protective effect was larger for families with lower level of education (Metayer et al., 2014).

2.1.4 Suggested associations

There are results from a recent meta-analysis that any vaccination before the age of one protects against childhood leukemia with an OR of 0.58 (95% CI 0.36, 0.91), however the nature of the finding is relatively imprecise as the effect could not be pinpointed to any specific vaccination (Morra et al., 2017). Allergies have also been studied as a potential etiological factor of childhood leukemia and a recent meta-analysis by Wallace et al. found an overall protective non-significant association related to childhood ALL (OR = 0.76, 95% CI 0.58, 1.01) (Wallace et al., 2018). When hay fever was analyzed separately, a significant protective odds ratio was observed (OR = 0.65, 95% 0.47, 0.90)

The effects of parental smoking were studied using CLIC data and no consistent effect was found with childhood AML for maternal smoking, before, during or after pregnancy. Only in subgroup analyses the authors observed an elevated risk among the Hispanic population and they conclude that the differing finding could be

explained by smoking habits or the underlying biology (Metayer et al., 2016). In the same study, paternal smoking was significantly associated with elevated risk of childhood leukemia in a pooled analysis (OR = 1.34, 95% CI 1.11, 1.62).

The effect of parental age was explored using CLIC data in a pooled analysis and increased paternal age was associated with childhood leukemia in analyses based on both self-reported survey data (OR = 1.05, 95% CI 1.00, 1.11) and register data (OR = 1.04, 95% CI 1.01, 1.07). A similar significant risk was associated with maternal age but only based on register-based studies (OR = 1.05, 95% CI 1.01, 1.08) (Petridou et al., 2018).

The effect of a family's socioeconomic status has also been studied and Kroll et al. reported based on English and Welsh register-based data higher childhood leukemia incidence in affluent communities (Kroll, Stiller, Murphy, & Carpenter, 2011). They hypothesized that this could be partly due to under-diagnosis of the lower categories or the higher SES categories could also be associated with other risk factors of childhood leukemia such as population mixing or delayed infection, which might be the more likely reason of the two (Greaves, 2006; Kinlen, 2012).

Parental exposure to pesticides was evaluated with a meta-analysis by Wigle et al. and they did not observe an overall association with paternal occupational pesticide exposure and childhood leukemia (OR = 1.09, 95% CI 0.88, 1.34) but for maternal prenatal occupational exposure they reported an OR of 2.09 (95% CI 1.51, 2.88) (Wigle, Turner, & Krewski, 2009). During the same year Turner et al. conducted a meta-analysis of the residential pesticide exposure and they report an overall OR of 1.54 (95% CI 1.11, 2.11) for any pesticide exposure during pregnancy (Turner, Wigle, & Krewski, 2010). They reported significantly elevated associations also for insecticide and herbicide exposures during pregnancy as well as for pesticide and insecticide exposures during childhood. These analyses were, however, based on self-reported exposures and are, thus, prone to responder and recall bias. A more recent pooled analysis reported an elevated OR for maternal pesticide exposure related to AML (OR = 1.94, 95% CI 1.19, 3.18) and paternal pesticide exposure related to ALL (OR = 1.20, 95% CI 1.06, 1.38) (Bailey et al., 2014). The paternal exposure appeared to be driven by T-ALL patients over the age of five years.

The effect of non-ionizing electromagnetic radiation has been studied and the International Agency for Research on Cancer (IARC) concluded in 2002 that extremely low frequency magnetic fields belong to a category of possible human carcinogens based on studies on childhood leukemia. A review by Schuz et al. concludes that there is insufficient evidence to alter this conclusion (Schuz & Ahlbom, 2008). The authors also summarize that the attributable fraction of

extremely low frequency magnetic field exposure is low (1 – 4%) if causality is assumed. A more recent pooled analysis concludes similarly that magnetic field should remain a possible carcinogen. For the highest exposure category of >0.3 μT they observed an OR of 1.44 (0.88, 2.36) (Kheifets et al., 2010).

Spatio-temporal clustering of childhood leukemia has been studied and in a recent Swiss study the authors observed clustering related especially to TEL-AML1 genetic subtype (Kreis et al., 2017). The OR for a TEL-AML1 positive presenting from a cluster was 2.54 (95% CI 1.52, 4.23) when compared to non-clustered cases. There are also earlier studies with results supporting a degree of clustering (Alexander et al., 1998; McNally, Alexander, & Bithell, 2006).

The reproductive histories of mothers have been analyzed in a case-control study by Ross et al. and overall, they reported a significant odds ratio related to previous miscarriages (OR = 1.45, 95% 1.00, 2.09) (Ross et al., 1997). In addition, there are earlier publications with concordant results (Kaye et al., 1991; Yeazel, Buckley, Woods, Ruccione, & Robison, 1995). Karalexi et al. published a meta-analysis on the same subject and reported an OR of 1.10 (95% CI 1.04, 1.18) for any fetal loss (Karalexi, Dessypris, Skalkidou, et al., 2017).

Marcotte et al. studied the effects of caesarean delivery on childhood leukemia using the compatible subset of CLIC data and observed that prelabour caesarean delivery was significantly associated with childhood leukemia in a large pooled analysis (1.23, 95% CI 1.04, 1.47). The risk was not associated with emergency caesarean sections and the authors hypothesized that the risk could be due to differing activation of the immune system in the absence of the normal stress response related to vaginal delivery. However, the authors suggest that future studies should try to account for the indication of the caesarean section as missing this information could have caused some bias to their own results. (Marcotte et al., 2016)

The effect of parental alcohol consumption was reviewed, and a meta-analysis of the existing studies was conducted by Karalexi et al. reporting an elevated dose-response association for AML (OR = 1.64, 95% CI 1.23, 2.17) when comparing moderate alcohol consumption to no consumption (Karalexi, Dessypris, Thomopoulos, et al., 2017). No association was observed for ALL in the meta-analysis. There are also highly putative results from a single Chinese case-control study that consumption of cured/smoked meat more than once per week could be related to higher risk of childhood leukemia (OR = 1.74, 95% CI 1.15, 2.64) whereas consumption of vegetables was observed to be associated with lower risks (OR = 0.55, 95% CI 0.37, 0.83) (the Kaohsiung Leukemia Research Group et al., 2009).

It has been hypothesized that MLL positive AML can arise after chemotherapy with topoisomerase II inhibitors. Consumed food also contains chemicals with this activity and thus maternal consumption of such foods could increase the risk of MLL positive AML of infants. This was investigated by Spector et al. and they observed that consuming foods with higher topoisomerase II inhibitor activity seemed to increase leukemia risk (OR = 3.2, 95% CI 0.9, 11.9 when comparing the lowest quartile to the highest). They also reported that consumption of fresh vegetables and fruits during pregnancy was associated with a decreased risk of childhood leukemia, particularly the MLL+ type (OR = 0.6, 95% CI 0.3, 1.1 when comparing the lowest quartile to the highest) (Spector, 2005).

Few studies have reported results suggestive of a degree of seasonality in the incidence of childhood leukemia pointing towards infectious etiology. Westerbeek et al. reported a significant excess of diagnoses during summer using data from the United Kingdom (Westerbeek et al., 1998). Based on analyses with data from the USA, Ross et al. were able to verify a higher incidence of ALL during summers (Ross et al., 1999).

2.2 Completeness of residential histories

When estimating environmental factors, typically, the location is the main determinant of exposure but changes in exposure levels occur over time as well. Domestic exposures are the most important source as children spend most of their time home in Finland (Mäkeläinen et al., 2005). The effect of complete residential histories in studies of background radiation and childhood leukemia is one possibility to explore the subject.

Kendall et al. reported results related to residential histories from their large British case-control study of childhood leukemia (Kendall, Wakeford, Bunch, Vincent, & Little, 2015). They found that 44% of cases had not moved between their birth and diagnosis. When they looked at only leukemia diagnoses under the age of five years, 55% of the cases had moved and 74% were living within 2 km of their original address. They observed that the mean gamma dose rate correlated well (0.90) between the county district of birth and county district of diagnosis.

Based on the same cohort used for the Swiss background radiation study, Kreis et al. observed that 34% of their cases had moved between birth and diagnosis (Kreis et al., 2017). A sub cohort of the Swiss data with stable residencies throughout the study period was analyzed separately and the HR for this cohort was found to be

1.046 (95% CI 0.999, 1.096) while the main result using the full data was 1.036 (95% CI 0.997, 1.077) (Spycher et al., 2015).

In the French study of background radiation and childhood leukemia, the authors note that 66% of the subjects had been living in the same municipality from birth to diagnosis (Demoury et al., 2017). They also reported a correlation coefficient of 0.89 for gamma radiation between the two addresses. The Danish study on indoor radon by Raaschou-Nielsen et al. noted that 58% of their subjects had lived in the same single-family house during the whole study period (Raaschou-Nielsen et al., 2008).

3 AIMS OF THE STUDY

The aims of the dissertation were:

1. To study the effect of low-dose ionizing radiation on childhood leukemia by investigating exposures from terrestrial background radiation, Chernobyl fallout and computerized tomography by using a case-control design (I, III).
2. To study the effect of having register-based complete residential histories on studies of environmental exposures (II)
   2.1. To evaluate the variation of natural background radiation doses in successive dwellings (II)
   2.2. To investigate the degree of residential mobility in Finnish families with young children (II)
3. To develop a statistical model for estimating indoor radon concentrations inside Finnish dwellings and to estimate the risk of childhood leukemia in relation to exposure to indoor radon gas (IV)

4 SUBJECTS AND METHODS

4.1 Finnish Register-based Case-control Study of Childhood Leukemia - FRECCLE

4.1.1 Case identification

Like some modern countries, Finland has a nationwide cancer registry (Finnish Cancer Registry). It covers all malignant diseases diagnosed in Finland after 1953 (Leinonen, Rantanen, Pitkäniemi, & Malila, 2016). Hospitals, physicians and laboratories are required by law (Act on National Personal Records Kept under the Health Care System, 556/1989) to notify the Finnish Cancer registry of every new case of childhood leukemia in addition to other malignant diseases without any consent of the patient. The Finnish Cancer registry also retrieves information on death certificates from Statistics Finland annually. In the case of missing clinical notification, the treating hospital is reminded. The coverage of Finnish Cancer Registry is estimated to be 86% for non-solid tumors for all population (Leinonen, Miettinen, Heikkinen, Pitkäniemi, & Malila, 2017). For pediatric (0–14 years) non-solid tumors the coverage has recently been estimated to be 97% (95% CI 89%, 94%) (Jokela, Leinonen, Malila, Taskinen, Madanat-Harjuoja, 2019). We identified all 0–14.99-year-old children diagnosed with leukemia between years 1990 and 2011. The diagnostic codes used in the collection were M9800 – M9948 according to ICD-O-3 classification. These diagnoses were further classified into five categories using diagnostic data from Finnish university hospitals: pre-B-ALL, T-cell acute lymphoblastic leukemia (T-ALL), unclassified ALL, AML and others. From Finnish Cancer Registry, we received data on the month of diagnosis and whether the child had a previous cancer history. In total, we received 1100 childhood leukemia cases.

4.1.2 Control selection

The Population Register Centre has essentially complete data on all Finnish citizens. We randomly selected three controls for every case. They were individually matched by the year of birth and sex to form case-control sets of four subjects (one case and three controls). To avoid correlation between the values of geological measurements, we opted not to match by geographical area. However, this could result in dissimilarities in the well-established risk factors inside case-control sets and could result in confounding. Four of the cases had prohibited the use of their information in Population Register Centre and three had invalid personal identification number in the Finnish Cancer Registry's database. After these exclusions, we had 4372 (1093 cases and 3297 controls) study subjects.

From the Population Register Centre we also collected the identities of the parents of the study subjects to be further linked to other databases for potential confounders and exposure data. We also collected data on the residential histories of the study subjects from the Population Register Centre.

We gained data on every start and end date of each residential period allowing us to map the complete residential history for each study subject. As some subjects were diagnosed in 1990 at 14 years of age, we needed data on the subjects starting from 1975. Each building in the subjects' residential histories was further linked with the Population Register Centre's building registry using the buildings' identification codes. The Building registry contained information on the type of building (house vs. apartment), building materials, area and volume of the building, number of floors, existence of basement, year of completion, type of ventilation and spatial coordinates in ETRS-TM35FIN geodetic datum.

For residential periods abroad (n=63, 0.8% of all residential periods), we had no accurate information and for some residential periods, only the municipality of the residency was known, and the exact dwelling could not be specified.

4.1.3 Exclusions

The main analyses were carried out using a latency period of two years. Differing time spans were tested in sensitivity analyses. With the study on background radiation, we modelled the exposure *in utero* by assuming an extra exposure period based on the gestational time in the dwelling of birth.

The use of a latency period effectively resulted in null exposure for subjects under the age of two at diagnosis. Apart from this, only seven cases were lost due to data

handling prohibitions or faulty identification codes. In the study on CT scans, we excluded patients with Down syndrome or with previous malignancies to control for confounding by indication. Subjects with major congenital malformations (heart, GI tract) were also excluded from in that study.

The length of the latency period (two years) was based on earlier publications. For example, in the study on CT scans this helps to avoid reverse causation. Also, the use of a latency period helps to eliminate passive doses, which in turn could result in underestimation of the studied potential risk factor. A passive dose is defined as the radiation dose received before the outcome of interest temporally too close to it to be able to contribute to its manifestation. The basic exclusions related to all published articles are summarized in the Figure 4.

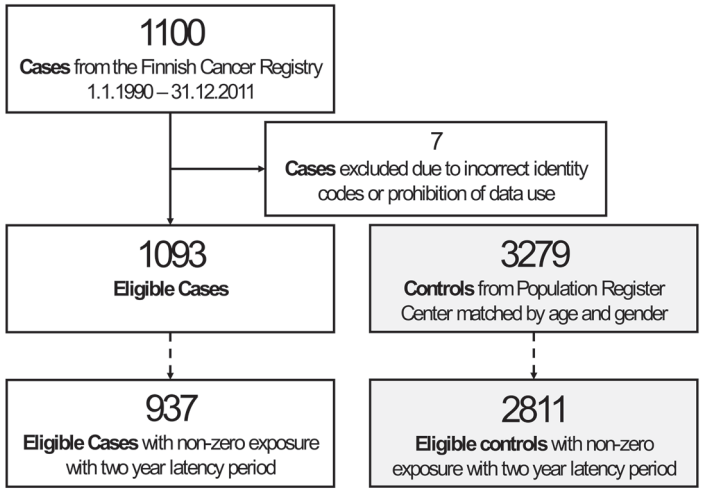


Figure 4. Flow-chart indicating identification of the cases and controls.

4.2 Data on exposures and explanatory variables

4.2.1 Congenital malformations

We collected data on congenital malformations from the National Institute for Health and Welfare's Congenital Malformation Register to be used as explanatory variables in the analyses. The register was established in 1962 and the data is sourced

from physicians, hospitals, prenatal and child-welfare clinics, cytologic laboratories, the Medical Birth Register, the Care Register for Health Care, the Register of Induced Abortions and the Register of Visual Impairment ("Register of Congenital malformations - THL," 2015). The register is also supplemented with data from Statistics Finland and the National Supervisory Authority for Welfare and Health. Information on the mother, pregnancy, infant, congenital anomalies and anomalies in family members are stored. The data on other malformations included data on Down syndrome but also on other malformations, possibly indicative of an underlying genetic syndrome.

4.2.2 Pregnancy and birth data

From the National Institute for Health and Welfare's Medical Birth Register we obtained data on the gestational weeks and birth weight. This register was established in 1987 and major reforms were carried out in 1990, 1996 and 2004 ("Medical Birth Register - THL," 2016). The register contains data on the mother, the pregnancy, the delivery, the infant, the infant's status at discharge and diagnoses up to the first week of life.

From the received information, we defined the "large for gestational age" variable as the values exceeding 90th percentile of birth weight relative to gestational age after assuming and verifying normality of the distribution. We also obtained data on maternal smoking during pregnancy, mode of delivery and parity. For maternal smoking during pregnancy the proportion of missing values was higher in the early study period but after the year 2000 the data was nearly complete and in total 16% of the data was missing.

4.2.3 Hospitalization periods

For cases and controls we collected data on all hospitalization periods. For each stay in the hospital, we obtained the start and end dates of the visit and the primary and secondary diagnoses assigned by the responsible physician. The register in its recent form was established in 1994 when it replaced the Hospital Discharge Register ("Care Register for Health Care - THL," 2016). The register contains data on patients discharged from inpatient care, day surgeries and specialized outpatient care. We received data on outpatient visits only for the last year of the study period (2011).

The data was utilized to look for diagnoses that could lead to confounding by indication related to computerized tomography scans.

4.2.4 Parental socioeconomic status and education level

From Statistics Finland, we collected data on parental socioeconomic status and education. The recent classification was established in 1989 and it is based on the recommendations by the United Nations. It contains seven classes based on social and economic characteristics ("Tilastokeskus - Luokitukset - Classification of Socio-economic Groups 1989 - Metadata," 2019). We received the data for both paternal and maternal socioeconomic status and education separately. Education was categorized into three classes (upper secondary, bachelor's degree and master's or doctor's degree). The socioeconomic status, on the other hand, had five classes (self-employed, upper level employees, lower level employees, manual workers and others). We received longitudinal data on five time points (1990, 1995, 2000, 2005, 2010) with five-year intervals. The data on both education and socioeconomic status were included in the models as potential confounders.

4.2.5 Computerized tomography scans

We collected data on all electronically available CT scans from ten of the largest hospitals in Finland. This included all five university hospitals in Finland and the five largest central hospitals. We estimated that this ten-hospital approach would cover approximately 90% of all pediatric CT scans during our study period which was deemed sufficient (III, supplementary material). Briefly, we extrapolated linearly the annual numbers based on data from Helsinki University Hospital with the assumption that central hospitals bought their first scanners after the university hospitals. The number of collected scans was further limited by the start date of the electronic record availability. For the university hospitals, all scans after year 1996 were available, but for the five central hospitals, all scans were available only after the year 2002. The data was recorded in the patient electronic medical records as procedure codes, which were used to derive all of the scan data used in the study. Some uncertainty is related to the procedure codes used as they changed multiple times during our study period and different hospitals adapted new codes at different times. This could lead to misclassification and some CT scans being missed in our database queries.

Eventually we decided to collect all available pediatric CT scans but prior to the direct collection of CT scans, we collected from the database of Tampere University Hospital also the diagnoses preceding the computerized tomography scans. We piloted an approach to identify indications that would lead to pediatric tomographies with high probability in search for a means to comprehensively approximate the number of CT scans in our study population. We queried all diagnoses 200 days before each pediatric computerized tomography scan and selected the most recent ones when more than one was found for a single patient. Also, thresholds of 100 and 50 days were experimented with. Eventually, no diagnoses highly indicative of a CT scan were found, and we preferred to directly collect all pediatric scans from the largest hospitals. However, decent likelihood ratios were observed for head traumas, other traumas and headaches but the absolute probabilities of a CT scan following the diagnoses remained low. The data on these experimental analyses will not be reported more specifically.

An experienced hospital physicist compiled a table of typical CT scan parameters used in pediatric CT scans during our study period. The data was obtained in a suitable format to be used with the National Cancer Institute Dosimetry system for Computed Tomography (NCICT) software. NCICT is a modern calculator for CT organ doses and it reports absorbed doses to organs based on patient (sex and age) and scan parameters (voltage, current, pitch, manufacturer, model, filters) (Lee, Kim, Bolch, Moroz, & Folio, 2015).

4.2.6 Geological data

To be used in the dose estimation models, we collected terrain elevation from sea-level from the Geological Survey of Finland as a high-resolution grid map (100m x 100m). We also collected data on the type of soil, which was further classified by its permeability by a three-tier variable. In addition, we collected data on the location of geological formations by the latest ice-age (mainly coarse gravel eskers). The maps of the soil type and geological formations were vector-based, and margins of the polygons were hand drawn by cartographers based on real geological measurement data (personal communication with employees of the Geological Survey of Finland). The vector maps were evaluated for reconstructing residential histories using GIS-based location scripts and basic R and VBA search algorithms. Both the elevation and soil type were used in the analyses of indoor radon. For the background radiation

study, the elevation was evaluated for cases and controls to approximate the potential variation in cosmic radiation.

4.2.7 Radiation data

We obtained 8 km x 8 km grid maps of the background radiation dose rates and soil's uranium concentration from the Radiation and Nuclear Safety Authority (Figure 5). A similar map on the Chernobyl fallout was also obtained together with longitudinal data on the decay of the fallout from fixed measuring stations to model the exponential decay and sinking of the radioactive matter. In Finland, the indoor measurements were carried out using thermoluminescent dosimeter chips and outdoor measurements using a mobile survey with a vehicle was conducted using a Geiger-Müller tube and a high-pressure ionization chamber (Arvela et al., 1995; Arvela, Markkanen, & Lemmelä, 1990). We also received municipality specific median background radiation and radon outdoor dose rates calculated by the Radiation and Nuclear Safety Authority.

We obtained the Ratikka database of indoor radon measurements, which consists of both representative survey measurements and privately purchased measurements. The database contains over 200 000 measurements in over 80 000 dwellings. The data contains the start and end dates of the measurement in addition to the address data with revised indoor radon concentrations. In Finland, indoor radon measurements are conducted using calibrated and standardized polycarbonate alpha-track detectors (Valmari et al., 2012).

An employee of the Radiation and Nuclear Safety Authority from the department of Radiation Practices Regulation compiled a list of all registered CT scanning devices in the ten largest Finnish hospitals. The list contained data on the manufacturer and the models of the scanners as well as the year of registration.

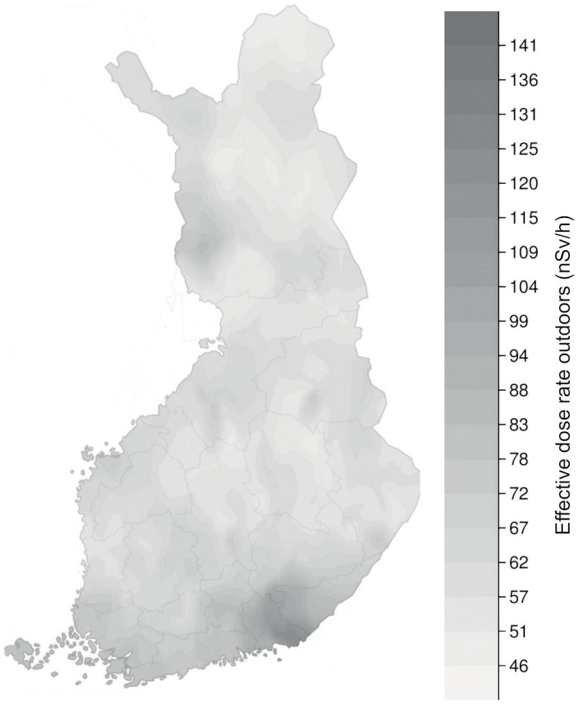


Figure 5. The effective background radiation dose rates in Finland outdoors

4.2.8 Genetic data

We collected electronically available more accurate diagnostic data on the diagnoses of childhood leukemia cases from databases of all five university hospitals. The genetic subtypes were accurately available for most cases but due to changes in the diagnostic processes during the study period and small numbers of some subtypes, we classified them into four categories based on consultation from an experienced clinician (TEL-AML1, High hyperdiploidy, other abnormalities and normal). The decision was also affected by the necessity to have sufficiently sized classes.

4.3 Dose estimation

4.3.1 Background radiation

The conversion from absorbed dose to equivalent RBM dose was achieved by first converting the absorbed dose rates in square maps to effective dose rates outdoors. A coefficient of 0.7 Sv/Gy was used to account for the shielding of tissues. Subsequently, the equivalent dose to red bone marrow was obtained using methods published by Kendall et al. by interpolating the conversion coefficients in between the three published point estimates for ages 8 weeks, 7 years and 18 years (Kendall et al., 2009; United Nations Scientific Committee on the Effects of Atomic Radiation, 2010). It was also necessary to account for the time spent indoors, i.e. occupancy, when estimating total doses. In general, a coefficient of 0.8 has been used to model occupancy (United Nations Scientific Committee on the Effects of Atomic Radiation, 2010). However, in Finland survey studies on children's age-specific occupancy have been conducted and this allows a more accurate approach (0-9 years: 0.75 and >9 years: 0.7) (Mäkeläinen et al., 2005).

Another aspect of dose estimation is taking into account the shielding properties of the different building materials of the dwellings. Rock-based building materials have been found to result in increased dose rates indoors. This is due to the fact that the radioactive elements of the building material are decaying. Also, the composition of the concrete typically reflects the composition of the local soil. In Finland, the outdoor dose rates can be converted into indoor dose rates in houses with a coefficient of 0.84 and a coefficient of 1.33 can be used for apartments. Houses are more commonly built from wood, while reinforced concrete is the most common material for apartments (Arvela et al., 1995).

Using these methods, we calculated the RBM doses related to each residential period and added them together to obtain the total cumulative dose from all residential periods of interest. Dividing the cumulative dose by the combined length of residential periods, we obtained a mean RBM dose rate to be used in the analyses.

4.3.2 Computerized tomography scans

Dosimetric data is stored in modern picture archiving and communication systems (PACS), used by hospitals. Hospitals manage their radiology departments with

radiological information systems (RIS) and Digital Imaging and Communications in Medicine (DICOM) is used as a standard image transfer service. In Finland, the use of databases allowing these functionalities started to become more common after the year 2000 (personal communication, Päivi Laarne) and since our study period extended to earlier years, we had to base our exposure assessment on different methods. We decided on using two distinct approaches when estimating exposure from pediatric computerized tomography scans. First, we used published tabulations of RBM doses from CT scans by Kim et al. (Kim et al., 2012). Their estimates take into account sex and age of the subject as well as the year when the scan was performed. These estimates were, however, based on data from the United Kingdom. Second, to improve on the estimates from the literature, we decided on using modern dose estimation software (NCICT) (Lee et al., 2015).

The NCICT software required data on the scanner manufacturer, scanner model, the body part, use of filters and the used scanning parameters (tube potential, current, pitch and collimation), which are crucial to obtaining an estimate of received dose. From the imaging codes and personal identification codes we derived the body part, imaging date and the age and sex of the subject. The data on registered scanners in Finland was obtained from the Radiation and Nuclear Safety Authority and we assumed that the scans were carried out using the most contemporary scanners available at each site in our main analyses. The scanning parameters for different age groups (0y, 1y, 5y, 10y and 15y) and different years (2002, 2004, 2006, 2008, 2010) were obtained from an experienced hospital physicist working at the Tampere University Hospital. After the RBM doses were calculated with NCICT, we modelled the use of contrast media and for examinations where it was utilized, we added an experimental extra RBM dose of 50% based on percentages published for other organs (Amato et al., 2013). The use of contrast media was derived from the scanning codes.

4.3.3 Indoor radon

To predict indoor radon concentrations for the residential periods of our study subjects, we constructed multiple statistical prediction models. When multiple measurements from a dwelling were available, we used the first one to minimize the effect of potential installations of radon protection solutions. The performance of the models was evaluated, and they were internally validated with a validation sample after building the model with the training sample. The models needed to be based

on variables that were nationally available for all Finns from the building database of the Population Register Center. This allowed them to be applicable for exposure assessment in nationwide studies. Before prediction, we used multiple imputations to complete the data on the predictors. After transforming the predicted indoor radon concentrations from a ten based logarithmic scale back to an arithmetic scale, we obtained geometric means for the predictions.

For subjects who had lived abroad, we used the world's average indoor radon concentration (39 Bq/m^3) as reported by UNSCEAR (United Nations Scientific Committee on the Effects of Atomic Radiation, 2010). For the subjects with data available only on municipality of residence, we used municipality specific averages received from Radiation and Nuclear Safety Authority.

For a small number of predictions (N=103, 1.4%), we had direct measurements available from the dwellings of study subjects for further validation of the model. Without considering the temporal aspect of the residential period and the measurement, we evaluated the performance of our predictions and reported coefficients of determination for the accuracy of predictions.

Using the predicted concentrations, we calculated cumulative exposures by multiplying the concentration by the length of the residential period. These cumulative exposures were summed together for each subject to obtain cumulative exposures for entire study periods. The average indoor radon concentration was also calculated for all study subjects by dividing the total cumulative exposure by the combined length of all residential periods.

We opted not to estimate RBM doses from the predicted radon exposures to ensure comparability to previous studies. Also, the variation in the RBM dose estimates calculated with published estimation methods is considerable and we had already introduced uncertainty in our exposure assessment by using prediction models instead of direct measurements (Harley & Robbins, 2009; Kendall & Smith, 2005).

4.4 Statistical methods

4.4.1 Conditional logistic regression

Regression analyses are a comprehensive set of methods designed to study the relationships between variables (Arora & Malhan, 2010). The basic idea is to have a

single outcome variable and one or multiple explanatory variables and the aim is to model a given change in the outcome variable as the change of explanatory variables. The explanatory variables can be, in general, binary, continuous or categorical variables. Using different link functions allows for different types of outcome variables (binary, ordered, categorical) under generalized linear models.

Regression analyses produce coefficients after fitting the model, which represent the association the relevant factor has with the outcome variable. For continuous variables, the fitted regression coefficient depicts a change in the outcome variable that a unit change on average a specific explanatory variable will produce. The models can be fit using a plethora of mathematical processes depending on the type of regression model. For example, an ordinary least squares method can be used with basic linear regression (Galton, 1886).

Logistic regression is a subtype of regression analyses, which is designed to model binary dependent variables by fitting a logistic function (Hosmer, Lemeshow, & Hosmer, 2013). Many extensions to logistic regression have been published (multinomial, ordered and conditional), which allow the use of different class variables as the dependent variable or a different configuration of study subjects.

With a binary outcome variable (leukemia, no leukemia) and individually matched study subjects, we used conditional logistic regression as the basis of our risk analyses. Both unadjusted and adjusted (multiple explanatory variables) analyses were performed. To control for confounding and for more accurate results, we included the available other explanatory variables into the models as potential confounders. However, the degree of missing data on the potential confounders widened our confidence intervals in some cases and when adjusting had no effect on the point estimate of interest, we preferred crude estimates.

In addition to producing odds ratios per unit increase of the exposure variable, we categorized the main explanatory variables into quartiles, calculated odds ratios for each of them and plotted the estimates with their confidence intervals. We used Bezier-smoothing when showcasing the background radiation results (Kim, Kim, Park, Hong, & Jeong, 1999). We also plotted linear dose-response curves for computerized tomography results on a logarithmic scale. Regarding the prediction models of residential radon concentrations, we experimented with quadratic explanatory variables in search of a model with a superior fit.

To avoid using the lowest quartile as the reference group with no estimation of uncertainty, we used floating absolute risk to gain narrower confidence intervals for the three other quartiles (Easton, Peto, & Babiker, 1991).

4.4.1.1 Odds ratio

Odds ratio is a widely used risk statistic (Cornfield, 1951). This is partly dictated by the fact that from case-control data it is not possible to derive relative risks and odds ratios are produced instead. Logistic regression is a major method that produces as its estimate odds ratios explaining its wide use in case-control studies. By its name, it is defined as a ratio of two odds. Odds, on the other hand, represent the likelihood of a certain event. Odds are defined as the number of positive binary events divided by the number of negative binary events. In comparison, a probability would be calculated with the total number of events as the denominator. In epidemiology, the relevant outcome is differentiating the two odds from each other and thus odds ratio is a means of estimating the effect of an exposure. According to the rare disease assumption, if the proportion of cases in the unexposed population approaches zero, an odds ratio starts to approach relative risk, which is more straight-forward to interpret. In many cases this applies to childhood leukemia as it is a rare disease. Odds ratios above unity suggest that the studied exposure increases the risk of the disease.

4.4.1.2 Covariate selection

Confounders are defined as variables with a connection to both the outcome at hand and the studied exposure. Also, the confounder cannot be a part of the causal chain of events leading to the outcome. If the frequencies of the confounding variable are not equal among differentially exposed study subjects, the coefficients estimated by regression analyses are confounded and do not represent the true association and this results in mixing of effects. As our study design is not randomized and we are not able to assume that both known and unknown confounders would be equally distributed we need to control for potential confounding. The preferable way to confront this hurdle is to adjust the used regression models with known and potential confounders so that the coefficients from the regression analyses are as reliable as possible. Another way to address the issue would be to stratify the analyses by categories of the potential confounder. Matching can also be utilized to avoid confounding but it should be limited to few strong and well-measured confounders to avoid other biases, such as sparse data bias (Mansournia, Jewell, & Greenland, 2018).

4.4.1.3 Evaluation of effect modification

In the case of positive association related to any given exposure, the effect might be driven by the whole study sample uniformly or there may be a smaller subgroup that is more responsible for the observed effect. For a subgroup, also an opposite, negative effect is possible. The complement sample might even be totally unaffected by the exposure.

To probe this effect, we ran subgroup analyses in which subsamples were analyzed separately and the coefficients could be compared to other subgroups. We also included multiplicative interaction terms in the regression models to represent the joint effect of the exposure of interest and the subgroup-defining variable. In the case of a statistically significant interaction term, we would conclude on effect modification by the grouping-variable. Effect modification can also be found between two independent variables, which both have their own individual effects, but their combined effect is not equal to the sum of the individual effects. However, the effect modifier is not required to have its own individual effect on the outcome.

4.4.2 Estimating indoor radon concentrations

4.4.2.1 Regression models

When building the prediction model for indoor radon concentrations, we opted to begin with a simple linear regression as the foundation. The indoor radon concentration was log-transformed to achieve normality and outliers were excluded if they differed by more than 3 standard deviations from the mean.

We also built models with categorical and binary dependent variables to gauge their performance and see whether the highest radon concentrations could be recognized. We categorized the indoor radon concentration into quartiles and used both ordered and multinomial logistic regression respectively. Also, a binary outcome variable with an 80% cutoff was modelled with traditional binomial logistic regression.

With the log-linear model we began modelling with all potential predictors and used a reverse selection algorithm to identify the predictors, which had the most favorable effect on Akaike's information criterion (AIC) (Akaike, 1973).

4.4.2.2 Multiple imputation

We used multiple imputation (MI) when building the radon prediction model to maximize the use of the available data (Honaker, King, & Blackwell, 2011). Using MI allowed us to use rows of data with missing data on some potential predictors. We constructed five imputed datasets with an R library Amelia and tested the log-linear models on each of them to ensure robust performance. Amelia uses a bootstrapping algorithm and based on the distributions of the known data, samples the values for the missing data producing a specified number of new datasets with no missing data for the analyses.

4.4.2.3 Model evaluation

We used the coefficient of determination, r^2, as the basis for comparing two models to each other after stepwise backwards predictor selection. It depicts the proportion of the outcome variable's variation that the explanatory variables are able to account for and spans from zero to one. We also preferred the adjusted r^2 over the crude one to adjust for the number of predictors (Peres-Neto, Legendre, Dray, & Borcard, 2006).

4.4.2.4 Machine learning methods

To complement traditional prediction methods, we experimented with more modern statistical methods. We wanted to validate the performance of the more traditional methods but also to find out whether better performance could be achieved with machine learning methods.

Random forests is an ensemble learning method that constructs numerous trees and in regression uses their mean (Breiman, 2001). It is able to internally correct its habit on significant overfitting by using sampled validation sets.

We also used deep neural networks with four defined layers (Schmidhuber, 2015). Deep neural networks are able to test for both linear and non-linear approximations and their performance was also evaluated with r^2.

4.4.3 Other statistical methods

We used generalized variation inflation factor (GVIF) to study the effect, the correlation between potential predictors in the radon models might have had on the obtained estimates (Fox & Monette, 1992). Nevertheless, the coefficients of determination reported for the models are accurate instead of correlations between predictor variables. However, the reported coefficients for specific predictors might become unstable due to intercorrelations.

We used Brant's test to evaluate if the assumption of parallel lines was met prior to performing ordered logistic regression analyses (Brant, 1990). The test is designed to differentiate between cases where one equation is not sufficient to depict relation between all levels of the outcome variable and multinomial logistic regression should be used instead.

Cohen's kappa was used to evaluate performance of the binary models. The test can be used to reliability between two raters and it is able to account for the expected degree of random variability (Cohen, 1960).

Spearman's rank-order correlation was used to represent statistical dependence between rankings of two variables (Spearman, 1904). Contrary to Pearson's correlation, it is also able to indicate non-linear associations. Both types of correlation span from zero to one, positive values indicate positive slopes and higher absolute values indicate stronger associations. However, neither type of correlation is able to measure the magnitude of the association.

4.4.4 Statistical software

The statistical analyses were performed using RStudio (Integrated Development for R. RStudio, Inc., Boston, MA). The R versions ranged from 3.2.0 to 3.5.3. Supportive and replicative analyses were ran using Stata 13.0, Excel 2016 and SAS 9.4.

4.5 Ethical considerations

Our study protocol was evaluated and supported by the Pirkanmaa Hospital District's Ethical Committee. All individual databases were contacted separately, and their specific permit protocols were followed. The data handling permits were obtained for all people participating in the project and handling the data. As the

studies in their whole were register-based and study subjects were not contacted, no informed consent was required in accordance with Finnish laws and regulations. This study was funded by the Väre foundation for childhood cancer research, the Pediatric Research Foundation, the Finnish Cultural Foundation and competitive state research funding of Pirkanmaa Hospital district (9T030 and 9U030).

5 RESULTS

5.1 Study population

Pre-B-ALL patients constituted the majority of diagnoses in our data (75.6%) and slightly over half of the cases were boys (52.0%) (Table 5). The median age at diagnosis of cases was 4.52 years (interquartile range (IQR) 2.72, 8.23).

Table 5. The characteristics of cases and controls before exclusions.

	Cases *(n=1093)*
Leukemia type	
pre-B-ALL	75.6% (826)
T-ALL	5.9% (64)
Unclassified ALL	1.8% (20)
AML	13.6% (149)
other	3.1% (34)
Age at diagnosis of the case (years)	
0 – 2	14.3% (156)
2 – 7	55.5% (605)
7 – 15	30.4% (332)
Gender	
Female	48.0% (525)
Male	52.0% (568)

The distributions of age and gender are similar for controls as they were used as matching variables.

Down syndrome was more common among the cases (OR = 60, 95% CI 14.5, 248) and no significant differences were observed between different education levels of the parents between cases and controls (Table 6). For socioeconomic status, we observed no differences between the different categories. Large for gestational age showed a statistically significant OR of 1.44 (95% CI 1.14, 1.81)

Table 6. Univariate conditional logistic regression models of the potential confounders

	Cases *(n=1093)*	**Controls** *(n=3279)*	**OR** (95% CI)
Large for gestational age			
No	86.7% (788)	90.1% (2493)	*reference*
Yes	13.3% (121)	9.9% (275)	1.44 (1.14, 1.81)
missing	184	511	*not applicable*
Mother's smoking during pregnancy			
No	83.1% (742)	84.5% (2296)	*reference*
Yes	16.9% (151)	15.5% (420)	1.15 (0.94, 1.42)
missing	200	563	*not applicable*
Down syndrome			
No	96.3% (1053)	99.9% (3277)	*reference*
Yes	3.7% (40)	0.1% (2)	60 (14.5, 248)
Parents' education			
Mother			
Upper secondary	48.5% (530)	50.6% (1659)	*reference*
Bachelor's degree	22.3% (244)	23.1% (756)	1.02 (0.84, 1.23)
Master's or doctor's degree	10.2% (112)	9.8% (321)	1.11 (0.87, 1.42)
missing	18.9% (207)	16.6% (543)	*not applicable*
Father			
Upper secondary	52.0% (568)	51.4% (1685)	*reference*
Bachelor's degree	15.2% (166)	16.2% (532)	1.09 (0.74, 1.14)
Master's or doctor's degree	10.0% (110)	10.2% (334)	0.98 (0.79, 1.31)
missing	22.8% (249)	22.2% (728)	*not applicable*
Parents' socioeconomic status			
Mother			
Self-employed	7.7% (84)	8.3% (273)	*reference*
Upper level employees	16.1% (176)	15.7% (514)	1.11 (0.83, 1.50)
Lower level employees	34.8% (380)	34.5% (1130)	1.09 (0.83, 1.44)
Manual workers	21.4% (231)	20.6% (674)	1.11 (0.83, 1.47)
Others	18.2% (199)	20.3% (664)	0.97 (0.72, 1.31)
missing	2.1% (23)	0.7% (24)	*not applicable*
Father			
Self-employed	13.9% (152)	12.0% (395)	*reference*
Upper level employees	17.6% (192)	18.2% (596)	0.85 (0.66, 1.08)
Lower level employees	18.3% (197)	17.9% (587)	0.87 (0.68, 1.12)
Manual workers	34.0% (372)	35.0% (1148)	0.86 (0.69, 1.07)
Others	12.4% (135)	14.3% (469)	0.75 (0.58, 0.98)
missing	4.1% (45)	2.6% (84)	*not applicable*

Large for gestational age was defined as the highest 10% of birth weights when adjusted for the duration of the pregnancy.

5.2 Background radiation and the risk of childhood leukemia (Publication I)

Overall, we observed a median RBM dose rate of 67.2 nSv/h for cases and 66.4 nSv/h for controls from terrestrial background radiation and Chernobyl fallout (Table 7). The exposure assessment is based on 8 x 8 km square maps of outdoor gamma radiation from Radiation and Nuclear Safety Authority. The median cumulative RBM doses were 1.96 mSv and 1.90 mSv, respectively.

Table 7. Cumulative doses and dose rates to RBM for cases and controls from natural background radiation and Chernobyl fallout

	Cases	Controls
Dose rate to RBM (nSv/h)		
Mean	69.9	69.4
Median	67.2	66.4
Interquartile range	56.3, 81.7	55.5, 81.9
Range	37.6, 154	33.7, 169
Cumulative dose to RBM (mSv)		
Mean	2.65	2.67
Median	1.96	1.90
Interquartile range	0.87, 4.06	0.83, 4.04
Range	0.01, 13.2	0.01, 14.8

The cumulative distribution functions showed no clear indication of a difference between the cases and the controls as the percentages on Y-axis for both cases and controls follow similar curves (Figure 6). From Chernobyl fallout, the median cumulative RBM doses were 0.1 mSv for both cases and controls. The median elevation of the cases' and controls' dwellings was 82 m (IQR 23 m, 108 m) and 74 m (IQR 21 m, 104 m), respectively.

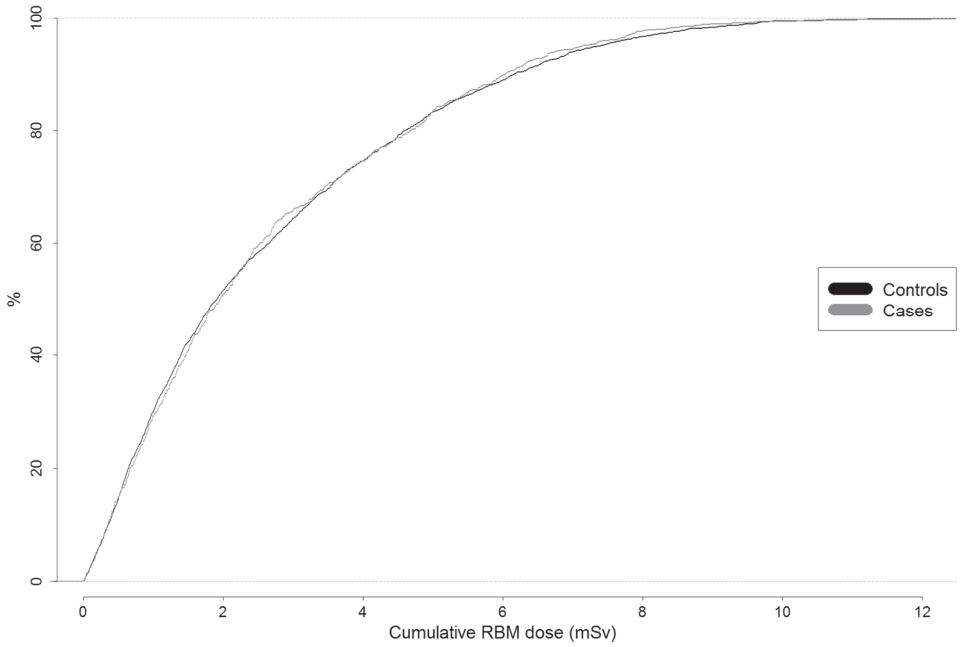


Figure 6. Cumulative distribution function of cumulative RBM dose from terrestrial background radiation and Chernobyl fallout

The OR of leukemia for a 10 nSv/h increase in average RBM dose rate was 1.01 (95% CI 0.97, 1.05) and for a 1 mSv increase in cumulative RBM dose we observed an OR of 0.97 (95% CI 0.89, 1.06). We also analyzed major histological subtypes alone and detected no difference between groups (ALL: OR = 1.02, 95% CI 0.98, 1.07, AML: 0.95, 95% CI 0.84, 1.08 for a 10 nSv/h increase in average RBM dose rate). However, we observed a significant (p=0.005, interaction) difference between the two age-groups with non-zero exposure (2–6.99 years and 7–14.99 years). The risk estimates were higher for the younger age-group (OR = 1.05, 95% CI 1.00, 1.10) and a lower point estimate was observed for the older age group (OR = 0.93, 95% CI 0.86, 1.00). The point estimates were higher for patients with high hyperdiploidy (OR = 1.11, 95% CI 1.02, 1.21) whereas the others genetic subtypes showed no statistically significant odds ratios. To explore the dose-response relation we divided the average RBM dose rate into quartiles and the medians of each quartile were respectively 48 nSv/h, 61 nSv/h, 71 nSv/h and 93 nSv/h and the odds ratios for each group using floating absolute risk were: 1 (95% CI 0.86, 1.16), 1.04 (95% CI

0.89, 1.21), 1.21 (95% CI 1.05, 1.39), 1.05 (95% CI 0.90, 1.22) showing a slight upward trend.

Exposure *in utero* was modelled by adding a 9-month exposure for each subject in their residence of birth and the results remained similar (OR = 1.01, 95% CI 0.97, 1.05 for a 10 nSv/h increase in average RBM dose rate). The analyses were also computed using no latency period and no difference in the results was observed (OR = 1.02, 95% CI 0.98, 1.05 for a 10 nSv/h increase in average RBM dose rate). The doses from Chernobyl fallout were negligible in comparison with the doses from terrestrial background radiation and we opted not to perform specific analyses on the subject as the average dose rates were so low.

We ran sensitivity analyses regarding subjects who had lived abroad and excluding them from the model did not yield different results (OR = 1.01, 95% CI 0.97, 1.06 for a 10 nSv/h increase in average RBM dose rate) nor did adjusting the model for residencies abroad. As another sensitivity analysis, the potential effect of CT scans was modelled using quantitative methods in multiple hypothetical bias scenarios: cases receiving 20% higher dose from CT scans than controls (OR = 1.03, 95% CI 0.99, 1.07 for a 10 nSv/h increase in average RBM dose rate), all subjects in the highest (OR = 1.01, 95% CI 0.97, 1.05 for a 10 nSv/h increase in average RBM dose rate) or the lowest quartile (OR = 1.01, 95% CI 0.97, 1.05 for a 10 nSv/h increase in average RBM dose rate) receiving 20% higher doses and only cases in the highest (OR = 1.02, 95% CI 0.98, 1.06 for a 10 nSv/h increase in average RBM dose rate) and the lowest quartile (OR = 1.02, 95% CI 0.98, 1.06 for a 10 nSv/h increase in average RBM dose rate) receiving a 20% higher dose than controls.

5.3 The effect of complete residential histories on studies of background radiation and childhood leukemia (Publication II)

Nearly half (48%) of the cases and controls had lived in only one address during their follow-up from diagnosis to reference date. The percentage of subjects who had lived in five or more dwellings during their study period was as low as five percent. The median of the number of dwellings for both cases and controls was 1.9 (range 1, 16 for cases and 1, 11 for controls). The data on the residency numbers followed a right-skewed log-normal distribution with 2% of the subjects having lived in six or more dwellings.

The median distance between successive dwellings for a family was under 4km (3.6 km, IQR 1.0 km, 12.1 km for cases and 3.3 km, IQR 0.9 km, 13.6 km for controls) and the median differences in dose rates were 2.4 nSv/h (IQR 0 nSv/h, 16.5 nSv/h) and 3.3 nSv/h (IQR 0 nSv/h, 17.7 nSv/h) for cases and controls respectively. The absolute differences were only slightly larger if the first and last addresses of the cases and controls were evaluated and the median separation between these dwellings was 4.3 km (IQR 1.2 km, 15.0 km) for cases and 4.2 km (IQR 1.2 km, 16.3 km) for controls as the differences in dose rate were 4.7 nSv/h (IQR 0 nSv/h, 17.0 nSv/h) and 3.8 nSv/h (0 nSv/h, 17.4 nSv/h) respectively.

The Pearson correlations between indoor RBM dose rates of two successive dwellings (0.62, 95% CI 0.60, 0.64) and the first and the last dwelling (0.67, 95% CI 0.65, 0.70) indicated moderate to strong correlation. The correlations between the outdoor dose rates were higher and the coefficients for the first and last dwelling were 0.91 (95% CI 0.89, 0.92) for cases and 0.88 (95% CI 0.87, 0.89) for controls.

When the conditional logistic regression models were fitted using the full residential histories which were available for all cases and controls (0.8%, N=63, of all residencies occurred abroad), we observed an OR of 1.01 (95% CI 0.97, 1.05) for an increase of 10 nSv/h in dose rate to the RBM. The odds ratios remained similar when the dose estimation was based only on the first and last address: the estimates were 1.02 (95% CI 0.99, 1.05) and 1.00 (95% CI 0.98, 1.03) for an increase of 10 nSv/h in the dose rate to the RBM, respectively. When we studied the subset of subjects who had lived in only one dwelling the OR was slightly higher at 1.05 (95% CI 0.98, 1.10) for an increase of 10 nSv/h in the dose rate to the RBM.

5.4 Computerized tomography scans and the risk of childhood leukemia (Publication III)

After excluding predisposing factors, there were eight exposed cases (12 scans) and nine exposed controls (12 scans) left in our material (Figure 7). Based on calculations with the NCICT we estimated the median RBM dose to be 10.1 mGy (IQR 4.8, 13.6) for exposed cases and 6.3 mGy (IQR 5.7, 7.1) for the exposed controls. The respective values based on available literature were 26.5 mGy and 17.6 mGy.

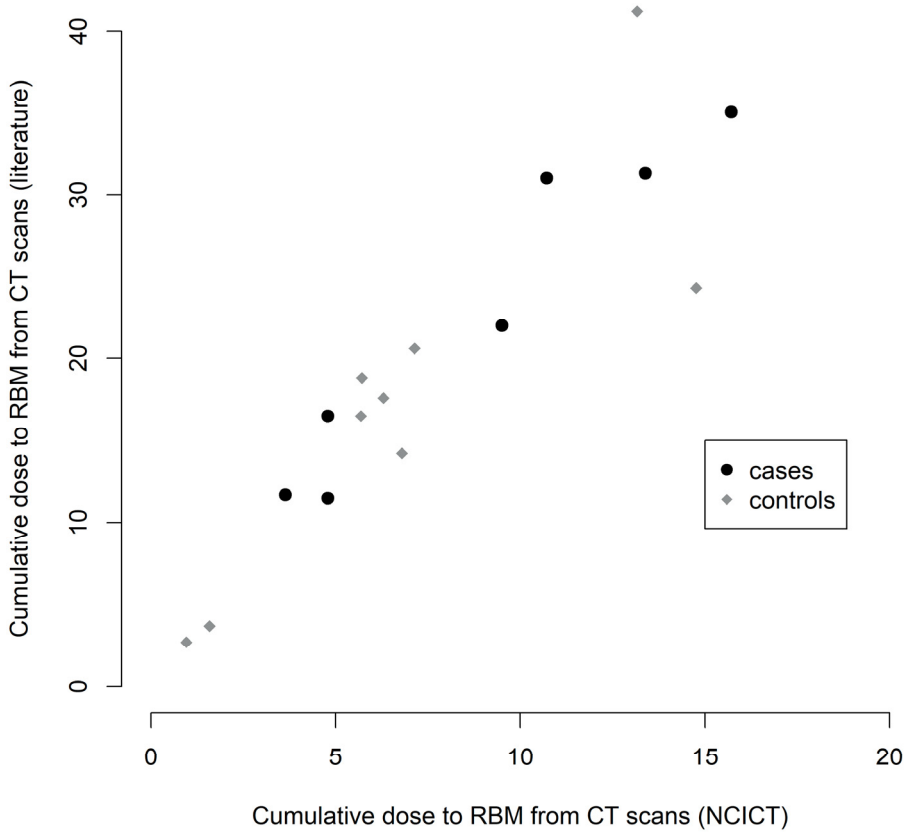


Figure 7. Scatterplot of cumulative RBM doses for the exposed study subjects based on the NCICT and literature

When subjects with one or more CT scans were compared with the unexposed sample, we observed an OR of 2.8 (95 % CI 1.1, 7.6) and when a similar analysis was restricted to subjects with two or more CT scans the OR rose to 6.22 (95% CI 0.89, 70). When subjects with one or more CT scans of the head were compared to the unexposed, the OR for childhood leukemia was 4.00 (95% CI 1.4, 11.5). The odds ratio for one or more CT scans was slightly higher for 2–6.99-year-old children (OR = 4.50, 95% CI 0.75, 26.9) when compared to 7–14.99-year-old children (OR = 2.27, 95% CI 0.68, 7.54) but the difference did not reach statistical significance ($p>0.1$).

We estimated the dose-response relationship and after fitting a linear model we observed an OR of 1.13 (1.02, 1.26) for an increase of 1 mGy in RBM dose. The equivalent estimate based on organ doses in the literature was slightly lower 1.05 (95% CI 1.01, 1.10) per 1 mGy of RBM dose. For pre-B-ALL the odds ratios were 1.14 (95% CI 1.02, 1.29) and 1.06 (1.01, 1.11) respectively.

When the cumulative exposure from one or more CT scans calculated with the NCICT was classified into tertiles and a fourth category consisting of subjects with no scans was added while the lowest (no CT scans) category was used as the reference category, the highest quartile was found to have an OR of 6.00 (95% CI 1.10, 32.7). The median doses for the three categories with scans were 3.9 mGy, 7.4 mGy and 13.7 mGy, respectively.

We performed explorative subgroup analyses and the OR for pre-B-ALL subjects was 1.14 (95% CI 1.02, 1.29) for an increase of 1 mGy based on literature-derived dose estimates. We also observed a slightly higher OR from one or more CT scans for subjects who were diagnosed at under seven years of age when compared to older children: 4.50 (95% CI 0.75, 26.9) vs. 2.27 (95% CI 0.68, 7.54), however the difference did not reach statistical significance (p=0.34).

We performed the analyses with Down syndrome patients only and observed a higher OR for leukemia when comparing those not exposed to CT scans with those DS subjects with one or more CT scans taken (OR = 5.21, 95% CI 2.19, 12.4). There was no evidence of differing risk by Down syndrome (p=0.99 for interaction).

As sensitivity analyses, we used the oldest available CT scanner at each site instead of the most modern. This resulted in slightly higher median doses among the exposed (9.7 mGy for cases and 7.1 mGy for controls). With these dose estimates we found an OR of 1.11 (95% CI 1.02, 1.22) for every 1 mGy increase in RBM dose. To control for a potential source of bias, we crudely explored whether subjects belonged unevenly to catchment populations of the largest ten hospitals and the chi-squared test performed showed a p-value of 0.30, suggesting no difference. We also included cumulative doses from background radiation in the models, but the original estimates remained robust.

5.5 Modelling residential radon and its effect on risk of childhood leukemia (Publication IV)

There were 93,219 measurements performed inside unique Finnish dwellings from recorded in the databases of the Radiation and Nuclear Safety Authority. Based on

these data, the median indoor radon concentration in Finland was 137 Bq/m^3 (IQR 68 Bq/m^3, 267 Bq/m^3). The distribution of indoor radon was log-normal with p99 reaching 1913 Bq/m^3. After filtering the data for better accuracy after database harmonization (concordance of the coordinates, concordance of the year of completion, missing critical data and duplicates), we had 73,903 measurements from houses and 3709 measurements from apartments for model development.

After creating the log-linear model with backwards selection, for houses, the largest proportion of variance in radon concentrations was explained by the year of completion (10.6%), while for apartments, it accounted for 4.6% of the variance. Another important variable was the soil permeability (3.0% for houses and 7.1% for apartments). The percentages are visually represented in Figure 8.

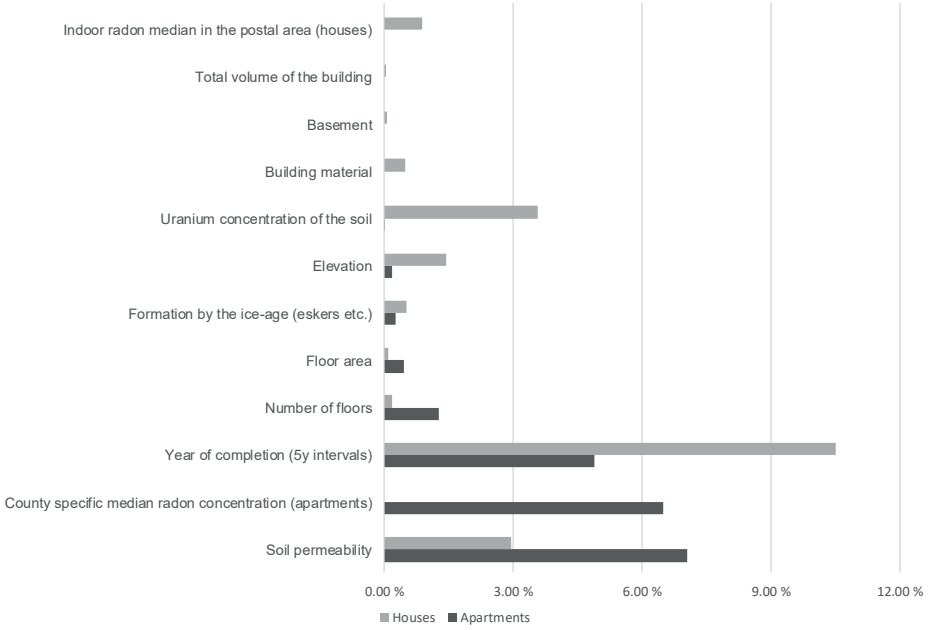


Figure 8. The proportions of variation in indoor radon concentration explained by different predictors using a log-linear prediction model

Based on the log-linear model, we created two categorical models. One based on quartiles of measured concentrations and another binary model with an 80-20 split. The quartile models reached a Cohen's kappa of 0.33 for houses and 0.38 for apartments. The kappas were lower for the binary model being 0.10 for houses and 0.25 for apartments. In exploratory analyses using multinomial and ordinal logistic regression, no major changes were observed, and for ordered logistic regression, the

assumption of parallel lines was not met for the categorized year of completion (Brant's test).

Using a traditional log-linear approach we reached a coefficient of determination of 0.21 for houses and 0.20 for apartments. Spearman correlations between the measured and predicted concentrations were 0.45 for houses and 0.44 for apartments. Using a random forests machine learning method, we obtained slightly higher estimates (0.28 for houses and 0.23 for apartments). With deep neural networks, we were not able to reach higher coefficients of determination (0.18 for houses and 0.19 for apartments). The models were not further calibrated manually to avoid losing the best fit found by the regression algorithm.

The Spearman correlation of predicted indoor radon concentrations was 0.45 with r^2 of 0.11 when the performance of the modelling was evaluated with available direct measurements. They were available for 1.4% of the residential periods when the residential periods of the study subjects were linked with Radiation and Nuclear Safety Authority's indoor radon database. The Pearson correlation between predicted cumulative exposures from log-linear and random forest models was 0.85.

The predicted cumulative exposures and time-weighted average concentrations are presented in the Table 8. Overall, indoor radon predictions based on random forest modelling were slightly higher.

Table 8. The medians and interquartile ranges of predicted indoor radon exposures with log-linear and random forests models for cases and controls

	median (IQR)	
	Cases	**Controls**
Log-linear		
Cumulative exposure (Bq/m³-years)	301 (121, 625)	292 (116, 636)
Average concentration (Bq/m³)	92 (68, 123)	89 (67, 121)
Random forests		
Cumulative exposure (Bq/m³-years)	357 (151, 789)	357 (152, 799)
Average concentration (Bq/m³)	107 (93, 127)	107 (93, 128)

In unadjusted analyses based on the log-linear model, higher indoor radon exposure showed no association with childhood leukemia (OR = 0.87, 95% CI 0.63, 1.19 for every 1000 Bq/m³-years increase in cumulative exposure). When the conditional logistic regression model was adjusted for the available potential confounders (large for gestational age, maternal smoking during pregnancy, Down syndrome, parental education and parental socioeconomic status), the OR rose to

1.06 (95% CI 0.59, 1.92), but did not reach significance. For the average concentration the crude and adjusted odds ratios were 0.99 (95% CI 0.97, 1.02) and 1.02 (0.99, 1.02) respectively for every 10 Bq/m^3 increase in average concentration. When the analyses were based on the random forests predictions, adjusting did not affect the odds ratios (unadjusted OR = 0.94, 95% CI 0.64, 1.37 for every 100Bq/m^3-years increase in cumulative exposure vs. adjusted OR = 0.93, 95% CI 0.42, 2.05 for every 100Bq/m^3-years increase in cumulative exposure).

We also explored relation in dose-response by plotting point estimates and their respective 95% confidence intervals for quartiles of average indoor radon concentration and a non-significant upward trend was observed based on both log-linear and random forests predictions for average concentration.

In exploratory subgroup analyses with adjusted models, we observed that the odds ratio for ALL patients based on the log-linear model was 1.32 (95% CI 0.67, 2.60) for every 1000 Bq/m^3-years increase in cumulative exposure. The younger age group (2–6.99 years) did not differ statistically from the older age group (7–14.99 years) ($p>0.1$).

5.6 Combined analysis

To study the effect of low-dose ionizing radiation as accurately as possible, we combined all available data on low-dose ionizing radiation exposure. In other words, we fit a model taking into account the studied exposures minimizing the potential bias related to the differential relations of the studied radiation sources (Table 9). The "combined gamma" category includes cumulative RBM exposure from background radiation from soil, building materials and Chernobyl fallout and the cumulative doses from pediatric CT scans. The analyses were adjusted for maternal smoking, LGA, parental education and parental socioeconomic status, and all subjects with Down syndrome were excluded. The OR for 1 mSv of RBM gamma dose was 1.05 (95% CI 0.96, 1.15) and for every 1000 Bq/m^3-years we observed an OR of 1.05 (95% CI 0.58, 1.90). In exploratory analyses, the pre-B ALL subgroup and the younger age group showed higher estimates, but the differences were not significant based on the interaction terms ($p>0.1$ for all analyses).

Table 9. The median exposures and odds ratios for combined gamma exposure and indoor radon exposure stratified by diagnostic and age categories.

	N		Median exposure (IQR)		OR (95% CI)
	Cases	Controls	Cases	Controls	
All subjects	913	2809			
Gamma combined (mSv)			1.93 (0.85, 4.12)	1.93 (0.84, 4.05)	1.05 (0.96, 1.15)
Indoor radon (1000 Bq/m^3-year)			0.29 (0.12, 0.65)	0.30 (0.13, 0.64)	1.05 (0.58, 1.90)
ALL	772	2377			
Gamma combined (mSv)			1.70 (0.79, 3.49)	1.76 (0.78, 3.73)	1.06 (0.93, 1.21)
Indoor radon (1000 Bq/m^3-year)			0.29 (0.11, 0.57)	0.27 (0.11, 0.58)	1.24 (0.62, 2.46)
Pre-B ALL	714	2203			
Gamma combined (mSv)			1.59 (0.74, 3.27)	1.65 (0.75, 3.46)	1.10 (0.95, 1.28)
Indoor radon (1000 Bq/m^3-year)			0.24 (0.10, 0.53)	0.25 (0.11, 0.55)	1.57 (0.74, 3.34)
AML	100	309			
Gamma combined (mSv)			3.48 (1.22, 5.92)	3.80 (1.22, 6.16)	1.04 (0.86, 1.25)
Indoor radon (1000 Bq/m^3-year)			0.54 (0.18, 0.94)	0.56 (0.20, 0.97)	0.37 (0.07, 1.95)
2–<7 years	581	1811			
Gamma combined (mSv)			1.05 (0.56, 1.78)	1.11 (0.53, 1.82)	1.35 (0.96, 1.89)
Indoor radon (1000 Bq/m^3-year)			0.16 (0.08, 0.28)	0.16 (0.08, 0.29)	2.02 (0.56, 7.37)
7–<15 years	332	998			
Gamma combined (mSv)			4.95 (3.96, 6.38)	4.87 (3.71, 6.37)	1.04 (0.94, 1.15)
Indoor radon (1000 Bq/m^3-year)			0.75 (0.54, 1.15)	0.81 (0.58, 1.14)	0.88 (0.44, 1.78)

6 DISCUSSION

6.1 Main findings

We estimated the childhood leukemia risk related to background gamma radiation, computerized tomography scans and indoor radon gas using a nationwide register-based data in case-control design in Finland. We also studied the effects of having complete residential histories in studies of childhood leukemia and the effects of background radiation on leukemia risk and constructed a statistical model for predicting indoor radon concentrations in Finland.

Overall, a small non-significant excess in risk of leukemia was observed for background radiation dose rate but the point estimate for cumulative dose was below unity. Regarding CT scans, we observed a significantly increased risk of childhood leukemia per unit of cumulative dose to RBM. However, the odds ratio was approximately an order of magnitude higher than was expected suggesting a potential presence of unknown bias. Higher indoor radon concentrations were related to higher odds ratios of childhood leukemia without reaching statistical significance. For the highest quartile of predicted radon exposure, we observed odds ratios between 1.1 and 1.3 depending on the modelling approach. Prediction models for indoor radon reached modest levels of accuracy and behaved robustly when internally validated. Due to nonideal performance of the prediction models the risk estimation contains a degree of uncertainty, which lowers the credibility of the results. The analyses on the residential histories showed that about half of the study subjects had not moved during the study period and those that had moved, lived often quite close to their previous dwelling (median 3.4km), suggesting that the lack of detailed residential histories does not pose an unsurmountable hurdle in risk assessment. However, we observed a higher odds ratio when subjects who had lived in only one dwelling were analyzed separately suggesting that studies relying only on the birth address might be overestimating the risks to some degree.

For the combined analysis of all separately studied exposures we observed an elevated yet non-significant odds ratio for combined gamma radiation from background radiation (terrestrial and Chernobyl fallout) and CT scans. As expected, the odds ratios were between the estimates from the separate studies of background

radiation and computerized tomography scans and the point estimates for cumulative exposure were in agreement with the risk magnitude expected based on the LSS study (Hsu et al., 2013). A risk estimate of the same magnitude, as in the separate study, was observed also for cumulative indoor radon exposure, but the confidence intervals were markedly wider than for the combined gamma exposure. In subgroup analyses, we observed slightly higher estimates for pre-B-ALL and the younger age-groups, but the differences were not significant.

6.2 Comparison to previously published results

6.2.1 Background radiation

Based on results published on Japanese life span studies the expected excess leukemia risk for a dose of 1 mSv would be approximately a RR of 1.01 with linear extrapolation for leukemia diagnoses other than adult T-cell leukemia and CLL for subjects under the age of 20 years (Hsu et al., 2013). It must be noted that the radiation dose was delivered rapidly and here also slowly accumulating exposures will be discussed. The results on background radiation were in line with the previous similar studies from United Kingdom and Switzerland and the confidence intervals overlapped markedly (Kendall et al., 2013; Spycher et al., 2015). In fact, the results published especially from United Kingdom are quite high in relation to the expected effect size. Our finding of higher risks for the younger age group is consistent with the results from the LSS, which has shown that risk varies significantly by age at exposure (Hsu et al., 2013). However, we observed unexpectedly low risk estimates for the older age group, which is probably driven by random error. There was no literature available to support the exploratory finding that subjects with high hyperdiploidy might have a higher risk estimate and the result needs to be validated in an independent dataset.

6.2.2 Computerized tomography scans

The main results of the CT scan study were higher than expected by almost an order of magnitude. When the exposure estimates were based on dose estimates from literature, we received an estimate much closer to the expected effect and the results were similar to the estimates reported from earlier large cohort studies (Mathews et

al., 2013; Pearce et al., 2012). The first study from the United Kingdom as well as the Australian study have been criticized for not accounting for potential predisposing factors, which might have biased the results to some extent (United Nations. Scientific Committee on the Effects of Atomic Radiation, 2013). In the case of reverse causation, symptoms from yet undiagnosed leukemia could lead to the CT imaging and this potential source of bias can be controlled for with latency periods of sufficient length. The fact that our results showed higher estimates for the scans of the head might be partly affected by predisposing factors that we could not account for as a study from the Netherlands found, in contrast to our study, a significant association with brain tumors but not with leukemia (Meulepas et al., 2018). The British study was able to combat the issue of predisposing factors by re-analyzing their results after collecting additional clinical information from radiology information systems and from pathology reports, which reduced their central estimate by 15% (Berrington De Gonzalez et al., 2016). We searched outpatient registries for potential predisposing factors and excluded subjects with Down syndrome and subjects with previous malignancies, which helped us avoid confounding by indication.

The average dose levels in our study (8.2 mGy) were comparable to studies from the Netherlands (9.5 mGy) and Germany (11.7 mGy) but the average dose to cases in the British study was considerably higher (18.5 mGy) (Krille et al., 2015; Meulepas et al., 2018; Pearce et al., 2012). For the French and Australian studies, the doses were on average 6.9 mGy and 4.6 mGy respectively (Journy et al., 2015; Mathews et al., 2013). Studies on X-ray examinations have also been linked to childhood leukemia without dose calculations with risk estimates below two but there are also studies that did not report significantly elevated risks (Bartley et al., 2010; Infante-Rivard, 2003; Meinert et al., 1999; Rajaraman et al., 2011). The inconclusive results regarding X-ray examinations are partly explained by the fact that the doses from CT scans are orders of magnitude greater than those from plain x-ray imaging which were not explored in this thesis.

6.2.3 Indoor radon

Our radon prediction model performed adequately and using random forests machine learning methods proved to be a helpful approach. However, we lacked the variables needed to predict the highest indoor radon concentrations, which could result in dilution of the effect if a real risk is assumed. Modelling the highest end of

the indoor radon distribution is not feasible as it is governed by an intricate set of coexisting attributes.

The Swiss modelling approach reached a r^2 no better than ours, but earlier attempts have been able to reach higher coefficients of determination (Andersen et al., 2007; Hauri et al., 2012). It has been suggested that the earlier models reached higher percentages of explained variance due to their smaller sample size but also the width of the indoor radon distribution plays a role. For example, the number of measurements the Danish model had, was approximately 5% of the number of measurements our model was based on. In addition, Finland has especially high variance in measured radon concentrations and the data on the important determinants of extreme values were not, unfortunately, available in nationwide registries. Predictors in the earlier models have included the same variables as our model did and after variable selection algorithms dwelling type, basement, floor, soil geology, region, building material, soil permeability, degree of urbanization and the year of construction remained in the model (Andersen et al., 2007; Hauri et al., 2012). Apart from the degree of urbanization, our models utilized the same variables.

We did not observe a significantly increased risk on childhood leukemia related to higher concentrations of indoor radon concentrations. After the Danish results were published, no study has been able to replicate the positive finding (Raaschou-Nielsen et al., 2008). Modelling-based studies from Norway, United Kingdom and Switzerland did not report any association with higher radon concentrations (Del Risco Kollerud et al., 2014; Hauri et al., 2013; Kendall et al., 2013). In combination with the previous results, our results confirm the notion that radon concentrations found in homes of families with children do not increase the risk of childhood leukemia at least to a noticeable degree. Complete residential histories were available only in the Danish study in addition to ours. In addition, our study might underestimate the magnitude of the possible association as high radon concentrations could not be reliably assigned.

Our findings are in concordance with older studies with direct measurements available, which are considerably more reliable in that domain (Kaletsch et al., 1999; Lubin et al., 1998; Steinbuch et al., 1999; UK Childhood Cancer Study Investigators, 2002b). A meta-analysis by Tong et al. found a significant positive association in ecological studies on childhood leukemia and indoor radon but in light of the multiple studies with direct measurements, the studies likely overestimate the true effect (Tong et al., 2012).

6.2.4 Residential histories

The evidence on the effects of having complete residential histories is limited and in the particular context of childhood leukemia the exact subject has not been studied as directly before. Earlier studies that have reported data on residential histories were in agreement regarding the proportion that approximately half of the children had moved during the study period (Demoury et al., 2017; Kendall et al., 2015; Kreis et al., 2017). The conclusion that meaningful results can be published also with analyses based on the dwelling of birth or diagnosis is also supported by the fact that the ORs reported for these analyses were similar. In addition, it is reasonable to propose that the address of birth would be especially important due to the early development of the infant's immune system but with full residential histories separate restricted analyses can be performed with this hypothesis in mind. The fact that we observed slightly higher ORs for the subgroup of subjects who had lived in only one dwelling might partly be explained by the fact that this subset is, overall, slightly younger. This is due to the fact that the number of dwellings tends to increase with age, and we observed highest ORs for the younger age-group. The same phenomenon was interestingly observed for the Danish study on indoor radon (Raaschou-Nielsen et al., 2008). In general, having limited information on residential histories results in a certain degree of exposure misclassification. When misclassification is non-differential among the cases and controls it biases the risk estimates towards null.

6.3 Methodological considerations

6.3.1 Subject selection and research material

Generally, the strength of scientific evidence varies between different study designs (Rothman, 2014). Conducting a double-blinded randomized controlled trial, which is often regarded as a golden standard, is not feasible in our setting. High-dose ionizing radiation is a well-established risk factor for childhood leukemia and it is not ethically sound to purposefully irradiate children in the light of the widely used ALARA (as low as reasonably achievable) principle. This can be derived using one of the principal precepts of bioethics. *Primum non nocere,* first do no harm.

As an experimental setup was not possible, in general, the second- choice would be to conduct a large prospective cohort study. This would avoid the typical problem

of case-control studies, namely potential selection bias related to unrepresentative controls. However, the annual incidence of childhood leukemia in Finland is so low that even with an extensive cohort, the number of diagnosed cases would remain so small that gaining sufficient statistical power to study with the expected effect sizes would be difficult and a cohort of millions would be required.

Therefore, conducting a large case-control study with nationwide and register-based data is the optimal option. In fact, in the case of a rare disease like childhood leukemia it reaches higher information efficiency and is, in that respect, superior to a cohort design. Also, for case-control studies, it is convenient to manage the time frame of exposures. In other words, electing this approach allowed us to gain a reasonable number of cases to study the expected risk sizes and it did not involve randomizing likely dangerous exposures for children. The register-based approach eliminated recall bias and nationwide sampling of controls from complete population registry effectively essentially eradicated selection bias even though frequency-based matching was not possible due to software challenges.

We had also evaluated the statistical power of our case-control design *a priori* with estimated dispersion of the exposures and expected effect sizes based on previously published studies. For conditional regression analyses of 1:3 groups with continuous explanatory variables and a 20% probability of a false negative result (beta = 0.2), analyses would be expected to detect odds ratios of 1.06 using the probable variation in background radiation cumulative RBM dose (Lachin, 2008). Increasing the number of controls would have resulted in only marginal gains in power based on our simulations.

Regardless, the statistical power, or more specifically the lack of it, was a challenge in our risk analyses. For background gamma radiation it had been estimated that approximately 7800 cases with five times as many controls would be needed to reach power of 80% with the dose rates in United Kingdom (Little, Wakeford, Lubin, & Kendall, 2010). However, the precision of the exposure assessment is another important aspect and our results could provide new insight on that front. Still, our results on smaller subgroups, such as AML, had wide confidence intervals due to small sample size. In addition, working with smaller statistical power increases the probability of chance findings and due to small expected risk sizes, even a small degree of confounding is capable of blurring the effect of interest. The problem of low statistical power can be expected to accumulate with non-differential misclassification of exposure and these phenomena amplify the effects of each other, but regression calibration can be used to combat the issue of measurement error.

To summarize, the register-based control selection helped us avoid selection bias and collecting all data from registries instead of surveys, allowed us to avoid recall bias. By using matched controls, we were able to control for the effect of two known well-measurable potential confounders, gender and age. The cases were obtained from the nationwide Finnish Cancer Registry, which is known for its high quality and we were able to include multiple known and potential confounders in the regression models. Also, with the genetic subtypes of the diagnoses collected from hospital databases we were also able to explore the most common genetic subgroups of ALL separately despite the fact that the classification was crude, and the data was not comprehensive in earlier years (i.e. before the year 2000).

6.3.2 Confounding

Regarding the study on CT scans, we were able to control for potential predisposing factors using multiple methods. We had data on congenital malformations, we excluded subjects with previous malignancies, we excluded patients with Down syndrome and searched the care register for other predisposing factors. In all studies, we used a latency period of two years to avoid reverse causation and to minimize passive exposures, which could dilute the effects. Including doses that are unable to contribute to the leukemogenesis might have diluted the risk estimates. On the other hand, this made us lose all subjects under two years old and we were not able to analyze infant leukemia separately. In the study on background radiation we performed sensitivity analyses with no lag period and, regarding infant leukemia, modelled the exposure *in utero* but the results remained largely unchanged.

For the missing information of potential confounders, we used multiple imputation which introduces a degree of random error to the related variables. This is expected to dilute risk estimates. Another possibility to deal with the missing data, which has been found helpful in studies of indoor radon, would have been to use mean imputation with the controls' average values but we preferred the more modern approach for its accuracy (Weinberg, Moledor, Umbach, & Sandler, 1996).

6.3.3 The quality of exposure assessment

We made great efforts to achieve the highest standards possible for our exposure assessment. We were able to improve on previous studies by using individual level data on the building properties and by modelling occupancy using coefficients published specifically for Finnish children. In a recent UNSCEAR report, the

accuracy of exposure assessment has been emphasized as an important determinant of the usefulness of the results in etiology studies alongside sufficient sample size (United Nations. Scientific Committee on the Effects of Atomic Radiation, 2013).

In rare diseases, international collaborative efforts enable collection of sufficient sample sizes to evaluate risks with low expected effect size and prevalence. However, without accurate exposure assessment, the studies are not able to add their full potential value due to misclassification of the exposure. In the case of non-differential misclassification in continuous explanatory variables, the risk estimates are diluted towards null. The used exposure assessment should be validated with independent direct measurements when possible. For a small proportion of subjects in our study on indoor radon, we were able to achieve this, but the results were subpar.

We were not able to cover every source of radiation and for example data on other medical examinations apart from CT studies is missing and thus our combined analysis of all available radiation sources did not cover completely the annual radiation of a typical Finnish child. Based on Radiation and Nuclear Safety Authority's estimates, we were able to cover roughly 70% of the annual dose (Muikku, Bly, Lahtinen, et al., 2014). Annual effective doses from cosmic radiation, internal exposure from eating and inhaling natural radioactive elements, other medical examinations and procedures using ionizing radiation sum up to approximately 0.7 mSv per person. However, it was our informed decision to omit cosmic radiation. Cosmic radiation levels in Finland are relatively stable and differential exposure for cases and controls was not to be expected based on our analyses on terrain elevation.

For indoor radon and background radiation we used approaches based on cumulative exposure and exposure rate. This is useful because the analyses of the latter contribute equally at all ages as analyses the cumulative exposure, by definition, are driven more by the older subjects. Analyzing the average dose rates, in a way, adjusts the exposure for age and is preferable in that sense. However, both approaches can be rationalized based on the assumptions on biological mechanisms of leukemogenesis. If the simple accumulation of radiation dose is assumed as the primary risk mechanism, regardless of the dynamics of accrual, then analyzing primarily cumulative exposures is justifiable.

Our approaches on exposure assessment of background radiation and indoor radon were based on high-quality measurements by the Radiation and Nuclear Safety Authority. We also had complete residential histories available, which improved the dose estimates for approximately half of our subjects who had moved during the

study period. We were able to use building characteristics in our estimations of background radiation and we also had national survey data on age-specific occupancy of Finnish children.

The CT scans were only collected from the ten largest hospitals and the collection was limited to electronically available data. According to our calculations, this approach was expected to reach a coverage of nearly 90% of all available CT scans but if the missed examinations were differentially distributed this would bias the results. However, with the latency period and control for predisposing factors, we expect this effect to remain minimal. Assuming that all missing scans (N=3) were performed to controls, the main results would remain unaffected. In addition, we evaluated also if the catchment areas of the chosen hospitals covered cases and controls equally and no evidence of differential catchment was found. Ideally, we would have collected accurate information on individual doses from the picture archiving systems of the hospitals, but the information was only available for the last years of our study period. In addition, we lacked data on the weight and height of the children and the exact scanning parameters were not available to us and thus more detailed organ dosimetry could not be performed. With these data, our dose estimates would have been more accurate. Also, the data on other radiological examinations and procedures using ionizing radiation would have improved the accuracy of our exposure assessment.

In the study on computerized tomography scans, we relied on expert opinion on the scanning parameters. This could have led to underestimation of the doses from the earlier scans leading to overestimation of our risk estimates. This could partly explain the higher than expected effect sizes. Also, the assumption to use the newest scanner at each site probably led to misclassification. However, the results remained robust, as in sensitivity analyses the oldest scanner was used. Both of these inaccuracies are most likely non-differential regarding the case/control status.

Regarding indoor radon, direct measurements are the preferred option when compared to prediction-based approaches. There exists a multitude of studies that faced challenges with residential radon predictions and among environmental exposures residential radon has proven to be especially difficult to predict. The dilemma remains between choosing more accurate measurements versus larger sample size and we preferred the latter. Regardless, our models reached r^2 values up to 0.28, which is higher than the 0.20 used by the Swiss study (Hauri et al., 2012).

We were able to properly evaluate our exposure assessment only for residential radon, for which the evaluation was the most needed, and the performance was found to be suboptimal. Even though variation in indoor radon measurements is

very high compared to terrestrial gamma radiation, the performance of the radon predictions underlines the importance of validation of background radiation predictions as well. Also, the radon measurement database of the Radiation and Nuclear Safety Authority is not totally representative of the distribution of indoor radon concentrations in Finland as dwellings with higher expected values were preferred in many cases to evaluate the need for protective measures.

The exposure assessment of individual study subjects was not directly based on real measurements in any of the studies and, for example, using geo-based estimations for background radiation will inevitably cause some misclassification, diluting the results. Also, the exposure assessment of indoor radon and CT scans included uncertainties. Even though we had data on many potential confounders, we were not able to include all currently known risk factors into our model due to limited data availability. We lacked data on maternal folate intake, paternal smoking, daycare attendance, breastfeeding, elective caesarean section. However, we were able to perform few quantitative bias analyses. These analyses revealed no material confounding and the results remained robust.

6.4 Recommendations for future research

We encourage emphasis in future studies to be put on the reliability of exposure assessment. An unavoidable requirement of this goal is the need to carry out onsite measurements at least to reliably verify exposure assessment methods. The optimal choice would be to base the analyses on direct measurements and settings allowing this approach should be favored. In addition, when uncertainty remains in exposure assessment, quantitative bias analyses should be utilized to quantify its effect on the results.

Another important aspect is to perform *a priori* calculations on statistical power, which is expected to guarantee sufficiently high statistical power (>80%). International collaborations are an obvious means for increasing statistical power in the case of a rare outcome and projects with shared funding and pre-specified authorship guidelines will provide a sustainable platform. The CLIC group is a good example of an effective collaboration and our group will continue participating in collaborative analyses in the future (Metayer et al., 2013).

Our own material is limited to patients diagnosed in 2011 and expanding it to the present day would offer a marked increase in statistical power. In addition, the degree of missing data in registries has decreased. In particular, we would be able to gain

much more accurate data on the genetic subtypes for the most recent years. Even DNA and RNA (ribonucleic acid) sequencing data is becoming more available as open access publishing and collaborative efforts are becoming a favored practice. Inclusion of comprehensive mutational data in the analyses of low-dose ionizing radiation and childhood leukemia would offer an interesting new measure of effect.

7 SUMMARY AND CONCLUSIONS

The role of ionizing radiation as a risk factor for childhood leukemia has been well-established for high doses and first reports of the association are almost 70 years old (Folley, Borges, & Yamawaki, 1952). This study aimed to estimate the magnitude of the risk low doses of ionizing radiation pose to childhood leukemia. This was achieved by constructing a nationwide register-based Finnish case-control dataset with matched and thus representative controls. In addition, the data on potential confounders was collected from nationwide registries. Separate analyses were performed on the effects of background radiation, pediatric computerized tomography scans and indoor radon concentrations. We observed significantly elevated risks from CT scans based on dose estimates both from literature and from modern dose estimation software. The risks were, however, slightly higher than expected.

We observed estimates suggestive of elevated risk from natural background radiation. The risk was significantly higher for younger children and cases with high hyperdiploidy showed also higher risk estimates. The results are in agreement with previously published results as our confidence intervals overlap with the results from earlier high-quality studies markedly and the point estimates are of the same magnitude. We also showed that in the setting of background radiation and childhood leukemia, useful results can be produced even with limited data on the residencies of the subjects, but complete residential histories should be prioritized when available.

We modelled indoor radon concentration and developed a model to predict indoor radon concentrations in Finnish dwellings. We were able to reach performance comparable with previous models, but our approach lacked the ability to recognize the dwellings with high radon concentrations. The risk of childhood leukemia rose with higher indoor radon exposure categories in adjusted analyses and in subgroup analyses pre-B-ALL diagnosed at a young age showed higher point estimates but overall, our results do not support the existence of a risk.

To summarize, our results support the notion that even low doses of ionizing radiation are a risk factor for childhood leukemia and the results are consistent with LNT extrapolation from higher doses. However, with the effect sizes reported in

this study, the population attributable fractions remain low. In Finland, new buildings are already required to follow strict radon protection regulations and the risks of pediatric CT scans have been acknowledged earlier by election of magnetic resonance imaging as the diagnostic imaging of choice in children and a clinical culture of generally requiring the benefits of the examination to outweigh the harms. Also, there exists no feasible way to completely protect children against terrestrial gamma radiation. Thus, based on our results, no immediate policy changes are warranted to counteract the risks. Technically, our results suggest a slightly higher risk from gamma radiation than the LNT models expected but more evidence is required to validate the existence of a potential difference. Our results on CT scans and background radiation can be included in the next international combined analyses, potentially by CLIC, based on which radiation protection guidelines can be updated to produce even more accurate risk estimates. Future studies should aim to improve on both statistical power and accuracy of the exposure assessment.

8 REFERENCES

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. *Proceedings of the 2nd International Symposium on Information Theory*, 267–281.

Alexander, F., Boyle, P., Carli, P.-M., Coebergh, J., Draper, G., Ekbom, A., … Vatten, L. (1998). Spatial temporal patterns in childhood leukaemia: further evidence for an infectious origin. *British Journal of Cancer*, *77*(5), 812–817. https://doi.org/10.1038/bjc.1998.132

Alter, B. P. (2014). Fanconi anemia and the development of leukemia. *Best Practice & Research Clinical Haematology*, *27*(3–4), 214–221. https://doi.org/10.1016/j.beha.2014.10.002

Amato, E., Salamone, I., Naso, S., Bottari, A., Gaeta, M., & Blandino, A. (2013). Can contrast media increase organ doses in ct examinations? a clinical study. *American Journal of Roentgenology*, *200*(6), 1288–1293. https://doi.org/10.2214/AJR.12.8958

Amitay, E. L., & Keinan-Boker, L. (2015). Breastfeeding and Childhood Leukemia Incidence: A Meta-analysis and Systematic Review. *JAMA Pediatrics*, *169*(6), e151025–e151025. https://doi.org/10.1001/jamapediatrics.2015.1025

Andersen, C. E., Raaschou-Nielsen, O., Andersen, H. P., Lind, M., Gravesen, P., Thomsen, B. L., & Ulbak, K. (2007). Prediction of 222Rn in danish dwellings using geology and house construction information from central databases. *Radiation Protection Dosimetry*, *123*(1), 83–94. https://doi.org/10.1093/rpd/ncl082

Arora, P. N., & Malhan, P. K. (2010). *Biostatistics*.

Arvela, H., Holmgren, O., & Reisbacka, H. (2012). Radon prevention in new construction in Finland: a nationwide sample survey in 2009. *Radiation Protection Dosimetry*, *148*(4), 465–474. https://doi.org/10.1093/rpd/ncr192

Arvela, H., Holmgren, O., Reisbacka, H., & Vinha, J. (2014). Review of low-energy construction, air tightness, ventilation strategies and indoor radon: results from Finnish houses and apartments. *Radiation Protection Dosimetry*, *162*(3), 351–363. https://doi.org/10.1093/rpd/nct278

Arvela, H., Hyvönen, H., Lemmelä, H., & Castrén, O. (1995). Indoor and Outdoor Gamma Radiation in Finland. *Radiation Protection Dosimetry*, *59*(1), 25–32. https://doi.org/10.1093/oxfordjournals.rpd.a082633

Arvela, H., Markkanen, M., & Lemmelä, H. (1990). Mobile Survey of Environmental Gamma Radiation and Fall-Out Levels in Finland After the Chernobyl Accident. *Radiation Protection Dosimetry*, *32*(3), 177–184. https://doi.org/10.1093/oxfordjournals.rpd.a080734

Aurengo, A., Averbeck, D., & Bonnin, A. (2005). Dose-effect relationships and estimation of the carcinogenic effects of low doses of ionizing radiation. *Academy of Sciences - National Academy of Medicine*, 58.

Auvinen, Hakama, M., Arvela, H., Hakulinen, T., Rahola, T., Suomela, M., … Rytomaa, T. (1994). Fallout from Chernobyl and incidence of childhood leukaemia in Finland, 1976-92. *BMJ*, *309*(6948), 151–154. https://doi.org/10.1136/bmj.309.6948.151

Auvinen, Seppä, K., Pasanen, K., Kurttio, P., Patama, T., Pukkala, E., … Hakulinen, T. (2014). Chernobyl fallout and cancer incidence in Finland 1988-2007: Chernobyl and Cancer in Finland. *International Journal of Cancer*, *134*(9), 2253–2263. https://doi.org/10.1002/ijc.28554

Azzam, E. I., Jay-Gerin, J.-P., & Pain, D. (2012). Ionizing radiation-induced metabolic oxidative stress and prolonged cell injury. *Cancer Letters*, *327*(1–2), 48–60. https://doi.org/10.1016/j.canlet.2011.12.012

Bailey, H. D., Fritschi, L., Infante-Rivard, C., Glass, D. C., Miligi, L., Dockerty, J. D., … Schüz, J. (2014). Parental occupational pesticide exposure and the risk of childhood leukemia in the offspring: Findings from the childhood leukemia international consortium: Parental Pesticides and Childhood Leukemia. *International Journal of Cancer*, *135*(9), 2157–2172. https://doi.org/10.1002/ijc.28854

Barros-Dios, J. M., Ruano-Ravina, A., Gastelu-Iturri, J., & Figueiras, A. (2007). Factors underlying residential radon concentration: Results from Galicia, Spain. *Environmental Research*, *103*(2). https://doi.org/10.1016/j.envres.2006.04.008

Bartley, K., Metayer, C., Selvin, S., Ducore, J., & Buffler, P. (2010). Diagnostic X-rays and risk of childhood leukaemia. *International Journal of Epidemiology*, *39*(6), 1628–1637. https://doi.org/10.1093/ije/dyq162

Bazyka, D., Prysyazhnyuk, A., Gudzenko, N., Dyagil, I., Belyi, D., Chumak, V., & Buzunov, V. (2018). Epidemiology of Late Health Effects in Ukrainian Chornobyl Cleanup Workers: *Health Physics*, *115*(1), 161–169. https://doi.org/10.1097/HP.0000000000000868

Belson, M., Kingsley, B., & Holmes, A. (2007). Risk Factors for Acute Leukemia in Children: A Review. *Environmental Health Perspectives*, *115*(1), 138–145. https://doi.org/10.1289/ehp.9023

Berrington De Gonzalez, A., Salotti, J. A., Mchugh, K., Little, M. P., Harbron, R. W., Lee, C., … Pearce, M. S. (2016). *Relationship between paediatric CT scans and subsequent risk of leukaemia and brain tumours: assessment of the impact of underlying conditions*. *114*. https://doi.org/10.1038/bjc.2015.415

Boice, J. D., Bigbee, W. L., Mumma, M. T., & Blot, W. J. (2003). Cancer incidence in municipalities near two former nuclear materials processing facilities in Pennsylvania. *Health Physics*, *85*(6), 678–690. https://doi.org/10.1097/00004032-200312000-00013

Boice, J. D., Mumma, M., Schweitzer, S., & Blot, W. J. (2003). Cancer mortality in a Texas county with prior uranium mining and milling activities, 1950–2001. *Journal of Radiological Protection*, *23*(3), 247–262. https://doi.org/10.1088/0952-4746/23/3/302

Boice, J. D., Mumma, M. T., & Blot, W. J. (2010). Cancer Incidence and Mortality in Populations Living Near Uranium Milling and Mining Operations in Grants, New Mexico, 1950–2004. *Radiation Research*, *174*(5), 624–636. https://doi.org/10.1667/RR2180.1

Boulton, F. (2019). Ionising radiation and childhood leukaemia revisited. *Medicine, Conflict and Survival*, 1–27. https://doi.org/10.1080/13623699.2019.1571684

Brant, R. (1990). Assessing Proportionality in the Proportional Odds Model for Ordinal Logistic Regression. *Biometrics*, *46*(4), 1171. https://doi.org/10.2307/2532457

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Brenner, D. J. (2007). Computed Tomography — An Increasing Source of Radiation Exposure. *The New England Journal of Medicine*, 8.

Buckley, J. D., Buckley, C. M., Breslow, N. E., Draper, G. J., Roberson, P. K., & Mack, T. M. (1996). Concordance for childhood cancer in twins. *Medical and Pediatric Oncology*, *26*(4), 223–229. https://doi.org/10.1002/(SICI)1096-911X(199604)26:4<223::AID-MPO1>3.0.CO;2-L

Buckley, J. D., Robison, L. L., Swotinsky, R., Garabrant, D. H., Manchester, P., Nesbit, M. E., … Hammond, G. D. (1989). *Occupational Exposures of Parents of Children with Acute Nonlymphocytic Leukemia: A Report from the Childrens Cancer Study Group1' 2*. 9.

Care Register for Health Care - THL. (2016). Retrieved May 10, 2019, from https://thl.fi/en/web/thlfi-en/statistics/information-on-statistics/register-descriptions/care-register-for-health-care

Carlos-Wallace, F. M., Zhang, L., Smith, M. T., Rader, G., & Steinmaus, C. (2016). Parental, In Utero, and Early-Life Exposure to Benzene and the Risk of Childhood Leukemia: A Meta-Analysis. *American Journal of Epidemiology*, *183*(1), 1–14. https://doi.org/10.1093/aje/kwv120

Caughey, R. W., & Michels, K. B. (2009). Birth weight and childhood leukemia: A meta-analysis and review of the current evidence. *International Journal of Cancer*, *124*(11), 2658–2670. https://doi.org/10.1002/ijc.24225

Clarke, R. T., Van den Bruel, A., Bankhead, C., Mitchell, C. D., Phillips, B., & Thompson, M. J. (2016). Clinical presentation of childhood leukaemia: a systematic review and meta-analysis. *Archives of Disease in Childhood*, *101*(10), 894–901. https://doi.org/10.1136/archdischild-2016-311251

Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20(1), 37–46. https://doi.org/10.1177/001316446002000104

Cornfield, J. (1951). A method of estimating comparative rates from clinical data; applications to cancer of the lung, breast, and cervix. *Journal of the National Cancer Institute*, *11*(6), 1269–1275.

Darby, S., Hill, D., Auvinen, A., Barros-Dios, J. M., Baysson, H., Bochicchio, F., … Doll, R. (2005). Radon in homes and risk of lung cancer: collaborative analysis of individual data from 13 European case-control studies. *BMJ (Clinical Research Ed.)*, *330*(7485), 223. https://doi.org/10.1136/bmj.38308.477650.63

Del Risco Kollerud, R., Blaasaas, K. G., & Claussen, B. (2014). Risk of leukaemia or cancer in the central nervous system among children living in an area with high indoor radon concentrations: results from a cohort study in Norway. *Br J Cancer*, *111*(7), 1413–1420. https://doi.org/10.1038/bjc.2014.400

Demoury, C., Marquant, F., Ielsch, G., Goujon, S., Debayle, C., Faure, L., … Clavel, J. (2017). Residential Exposure to Natural Background Radiation and Risk of Childhood Acute Leukemia in France, 1990–2009. *Environmental Health Perspectives*, *125*(4), 714–720. https://doi.org/10.1289/EHP296

Easton, D. F., Peto, J., & Babiker, A. G. A. G. (1991). Floating absolute risk: An alternative to relative risk in survival and case-control analysis avoiding an arbitrary reference group. *Statistics in Medicine*, *10*(7), 1025–1035. https://doi.org/10.1002/sim.4780100703

Eden, T. (2010). Aetiology of childhood leukaemia. *Cancer Treatment Reviews*, *36*(4), 286–297. https://doi.org/10.1016/j.ctrv.2010.02.004

Farber, S., Diamond, L., Mercer, R., Sylvester, R., & Wolff, J. (1948). Temporary remissions in acute leukemia in children produced by folic antagonist, 4-aminopteroyl-glutamic acid (Aminopterin). *New England Journal of Medicine*, *238*(23), 787–793.

Folley, J. H., Borges, W., & Yamawaki, T. (1952). Incidence of leukemia in survivors of the atomic bomb in Hiroshima and Nagasaki, Japan. *The American Journal of Medicine*, *13*(3), 311–321. https://doi.org/10.1016/0002-9343(52)90285-4

Ford, A. M., Ridge, S. A., Cabrera, M. E., Mahmoud, H., Steel, C. M., Chan, L. C., & Greaves, M. (1993). In utero rearrangements in the trithorax-related oncogene in infant leukaemias. *Nature*, *363*(6427), 358. https://doi.org/10.1038/363358a0

Fox, J., & Monette, G. (1992). Generalized Collinearity Diagnostics. *Journal of the American Statistical Association*, *87*(417), 178–183. https://doi.org/10.1080/01621459.1992.10475190

Friedman, D. L., Whitton, J., Leisenring, W., Mertens, A. C., Hammond, S., Stovall, M., … Neglia, J. P. (2010). Subsequent Neoplasms in 5-Year Survivors of Childhood Cancer: The Childhood Cancer Survivor Study. *JNCI: Journal of the National Cancer Institute*, *102*(14), 1083–1095. https://doi.org/10.1093/jnci/djq238

Gale, K. B., Ford, A. M., Repp, R., Borkhardt, A., Keller, C., Eden, O. B., & Greaves, M. F. (1997). Backtracking leukemia to birth: Identification of clonotypic gene fusion sequences in neonatal blood spots. *Proceedings of the National Academy of Sciences*, *94*(25), 13950–13954. https://doi.org/10.1073/pnas.94.25.13950

Galton, F. (1886). Regression Towards Mediocrity in Hereditary Stature. *The Journal of the Anthropological Institute of Great Britain and Ireland*, *15*, 246. https://doi.org/10.2307/2841583

Garzillo, C., Pugliese, M., Loffredo, F., & Quarto, M. (2017). Indoor radon exposure and lung cancer risk: a meta-analysis of case-control studies. *Translational Cancer Research*, *6*(S5), S934–S943. https://doi.org/10.21037/tcr.2017.05.42

Graham, S., Levin, M. L., Lilienfeld, A. M., Schuman, L. M., Gibson, R., Dowd, J. E., & Hempelmann, L. (1966). Preconception, intrauterine, and postnatal irradiation as related to leukemia (full text not available). *National Cancer Institute Monograph*, *19*, 347–371.

Greaves. (2006). Infection, immune responses and the aetiology of childhood leukaemia. *Nature Reviews Cancer*, *6*(3), 193–203. https://doi.org/10.1038/nrc1816

Greaves, M. (2005). In utero origins of childhood leukaemia. *Early Human Development*, *81*(1), 123–129. https://doi.org/10.1016/j.earlhumdev.2004.10.004

Greaves, M. (2018). A causal mechanism for childhood acute lymphoblastic leukaemia. *Nature Reviews Cancer*, *18*(8), 471–484. https://doi.org/10.1038/s41568-018-0015-6

Greaves, M. F., Maia, A. T., Wiemels, J. L., & Ford, A. M. (2003). Leukemia in twins: lessons in natural history. *Blood*, *102*(7), 2321–2333. https://doi.org/10.1182/blood-2002-12-3817

Greaves, M. F., & Wiemels, J. (2003). Origins of chromosome translocations in childhood leukaemia. *Nature Reviews Cancer*, *3*(9), 639–649. https://doi.org/10.1038/nrc1164

Grosche, B., Lackland, D., Mohr, L., Dunbar, J., Nicholas, J., Burkart, W., & Hoel, D. (1999). Leukaemia in the vicinity of two tritium-releasing nuclear facilities: a comparison of the Kruemmel Site, Germany, and the Savannah River Site, South Carolina, USA. *Journal of Radiological Protection*, *19*(3), 243–252. https://doi.org/10.1088/0952-4746/19/3/302

Gunz, F. W., & Atkinson, H. R. (1964). Medical Radiations and Leukaemia: A Retrospective Survey. *BMJ*, *1*(5380), 389–393. https://doi.org/10.1136/bmj.1.5380.389

Haddy, N., Cécile Le Deley, M., Samand, A., Diallo, I., Guérin, S., Guibout, C., … de Vathaire, F. (2006). Role of radiotherapy and chemotherapy in the risk of secondary

leukaemia after a solid tumour in childhood. *European Journal of Cancer*, *42*(16), 2757–2764. https://doi.org/10.1016/j.ejca.2006.05.034

Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of Cancer: The Next Generation. *Cell*, *144*(5), 646–674. https://doi.org/10.1016/j.cell.2011.02.013

Harley, N. H., & Robbins, E. S. (2009). Radon and Leukemia in the Danish Study: Another Source of Dose: *Health Physics*, *97*(4), 343–347. https://doi.org/10.1097/HP.0b013e3181ad8018

Harley, & Robbins, E. S. (1992). 222Rn Alpha Dose to Organs Other than Lung. *Radiation Protection Dosimetry*, *45*(1–4), 619–622. https://doi.org/10.1093/rpd/45.1-4.619

Hasle, H. (2001). Pattern of malignant disorders in individuals with Down's syndrome. *The Lancet Oncology*, *2*(7), 429–436.

Hauri, Huss, A., Zimmermann, F., Kuehni, C. E., & Röösli, M. (2012). A prediction model for assessing residential radon concentration in Switzerland. *Journal of Environmental Radioactivity*, *112*, 83–89. https://doi.org/10.1016/j.jenvrad.2012.03.014

Hauri, Spycher, B., Huss, A., Zimmermann, F., Grotzer, M., von der Weid, N., … Röösli, M. (2013). Domestic radon exposure and risk of childhood cancer: A prospective census-based cohort study. *Environmental Health Perspectives*, *121*(10), 1239–1244. https://doi.org/10.1289/ehp.1306500

Hawkins, M. M., Wilson, L. M., Stovall, M. A., Marsden, H. B., Potok, M. H., Kingston, J. E., & Chessells, J. M. (1992). Epipodophyllotoxins, alkylating agents, and radiation and risk of secondary leukaemia after childhood cancer. *BMJ*, *304*(6832), 951–958. https://doi.org/10.1136/bmj.304.6832.951

Henshaw, D. L., Eatough, J. P., & Richardson, R. B. (1990). Radon as a causative factor in induction of myeloid leukaemia and other cancers. *The Lancet*, *335*(8696), 1008–1012. https://doi.org/10.1016/0140-6736(90)91071-H

Honaker, J., King, G., & Blackwell, M. (2011). Amelia II: A Program for Missing Data. *Journal of Statistical Software*, *45*(7). https://doi.org/10.18637/jss.v045.i07

Hosmer, D. W., Lemeshow, S., & Hosmer, D. W. (2013). *Applied Logistic Regression*. Hoboken: John Wiley & Sons.

Hrušák, O., Trka, J., Zuna, J., Poloučková, A., Kalina, T., & Starý, J. (2002). Acute lymphoblastic leukemia incidence during socioeconomic transition: selective increase in children from 1 to 4 years. *Leukemia*, *16*(4), 720–725. https://doi.org/10.1038/sj/leu/2402422

Hsu, W.-L., Preston, D. L., Soda, M., Sugiyama, H., Funamoto, S., Kodama, K., … Mabuchi, K. (2013). The incidence of leukemia, lymphoma and multiple myeloma among atomic bomb survivors: 1950-2001. *Radiation Research*, *179*(3), 361–382. https://doi.org/10.1667/RR2892.1

Huang, W.-Y., Muo, C.-H., Lin, C.-Y., Jen, Y.-M., Yang, M.-H., Lin, J.-C., … Kao, C.-H. (2014). Paediatric head CT scan and subsequent risk of malignancy and benign brain tumour: a nation-wide population-based cohort study. *British Journal of Cancer*, *110*(9), 2354–2360. https://doi.org/10.1038/bjc.2014.103

Hutter, J. J. (2010). Childhood Leukemia. *Pediatrics in Review*, *31*(6), 234. https://doi.org/10.1542/pir.31-6-234

ICRP. (2007). ICRP 103. *Annals of the ICRP*, *37*(2–4), 1–332. https://doi.org/10.1016/j.icrp.2007.10.003

Inaba, H., Greaves, M., & Mullighan, C. G. (2013). Acute lymphoblastic leukaemia. *The Lancet*, *381*(9881), 1943–1955. https://doi.org/10.1016/S0140-6736(12)62187-4

Infante-Rivard, C. (2003). Diagnostic X-rays, DNA repair genes and childhood acute lymphoblastic leukemia: *Health Physics*, *85*(1), 60–64. https://doi.org/10.1097/00004032-200307000-00012

International Agency for Research on Cancer (Ed.). (2000). *X- and gamma-radiation, and neutrons*. Lyon: IARC.

Jablon, S., Hrubec, Z., & Boice, J. D. (1991). *Cancer in Populations Living Near Nuclear Facilities*. 6.

Jokela M, Leinonen MK, Malila N, Taskinen M, Madanat-Harjuoja LM. Completeness of pediatric cancer registration in the Finnish Cancer Registry. *Acta Oncologica*. 2019 Jul 10:1-4. doi: 10.1080/0284186X.2019.1638522.

Journy, N., Rehel, J.-L., Ducou Le Pointe, H., Lee, C., Brisse, H., Chateil, J.-F., … Bernier, M.-O. (2015). Are the studies on cancer risk from CT scans biased by indication? Elements of answer from a large-scale cohort study in France. *British Journal of Cancer*, *112*(1), 185–193. https://doi.org/10.1038/bjc.2014.526

Kaldor, J. M., Day, N. E., Clarke, E. A., Van Leeuwen, F. E., Henry-Amar, M., Fiorentino, M. V., … Assouline, D. (1990). Leukemia following Hodgkin's disease. *The New England Journal of Medicine*, *322*(1), 7–13. https://doi.org/10.1056/NEJM199001043220102

Kaletsch, U., Kaatsch, P., Meinert, R., Schüz, J., Czarwinski, R., & Michaelis, J. (1999). Childhood cancer and residential radon exposure - results of a population-based case-control study in Lower Saxony (Germany). *Radiation and Environmental Biophysics*, *38*(3), 211–215.

Karalexi, Dessypris, N., Skalkidou, A., Biniaris-Georgallis, S.-I., Kalogirou, E. I., Thomopoulos, T. P., … Petridou, E. Th. (2017). Maternal fetal loss history and increased acute leukemia subtype risk in subsequent offspring: a systematic review and meta-analysis. *Cancer Causes & Control*, *28*(6), 599–624. https://doi.org/10.1007/s10552-017-0890-2

Karalexi, Dessypris, N., Thomopoulos, T. P., Ntouvelis, E., Kantzanou, M., Diamantaras, A.-A., … Petridou, E. Th. (2017). Parental alcohol consumption and risk of leukemia in the offspring: a systematic review and meta-analysis. *European Journal of Cancer Prevention*, *26*(5), 433–441. https://doi.org/10.1097/CEJ.0000000000000350

Kaye, S. A., Robison, L. L., Smithson, W. A., Gunderson, P., King, F. L., & Neglia, J. P. (1991). Maternal reproductive history and birth characteristics in childhood acute lymphoblastic leukemia. *Cancer*, *68*(6), 1351–1355.

Kendall, G. M., Fell, T. P., & Harrison, J. D. (2009). Dose to red bone marrow of infants, children and adults from radiation of natural origin. *Journal of Radiological Protection: Official Journal of the Society for Radiological Protection*, *29*(2), 123–138. https://doi.org/10.1088/0952-4746/29/2/001

Kendall, G. M., Wakeford, R., Bunch, K. J., Vincent, T. J., & Little, M. P. (2015). Residential mobility and associated factors in relation to the assessment of exposure to naturally occurring radiation in studies of childhood cancer. *Journal of Radiological Protection*, *35*(4), 835–868. https://doi.org/10.1088/0952-4746/35/4/835

Kendall, Little, M. P., Wakeford, R., Kathryn, J., Miles, J. C. H., Vincent, T. J., … Murphy, M. F. G. (2013). A record-based case-control study of natural background radiation and the incidence of childhood leukaemia and other cancers in Great Britain during 1980 – 2006. *Leukemia*, *27*(1), 3–9. https://doi.org/10.1038/leu.2012.151.A

Kendall, & Smith, T. J. (2005). Doses from radon and its decay products to children. *Journal of Radiological Protection*, *25*(3), 241–256. https://doi.org/10.1088/0952-4746/25/3/002

Kheifets, L., Ahlbom, A., Crespi, C. M., Draper, G., Hagihara, J., Lowenthal, R. M., … Wunsch Filho, V. (2010). Pooled analysis of recent studies on magnetic fields and childhood leukaemia. *British Journal of Cancer*, *103*(7), 1128–1135. https://doi.org/10.1038/sj.bjc.6605838

Kim, Berrington de González, A., Pearce, M. S., Salotti, J. A., Parker, L., McHugh, K., … Lee, C. (2012). Development of a database of organ doses for paediatric and young adult CT scans in the United Kingdom. *Radiation Protection Dosimetry*, *150*(4), 415–426. https://doi.org/10.1093/rpd/ncr429

Kim, C., Kim, W., Park, B. U., Hong, C., & Jeong, M. (1999). Smoothing techniques via the bezier curve. *Communications in Statistics - Theory and Methods*, *28*(7), 1577–1597. https://doi.org/10.1080/03610929908832374

Kinlen, L. J. (2012). An examination, with a meta-analysis, of studies of childhood leukaemia in relation to population mixing. *British Journal of Cancer*, *107*(7), 1163–1168. https://doi.org/10.1038/bjc.2012.402

Kollerud, R., Blaasaas, K., Venvik, G., K. Daviknes, H., Aune, E., & Claussen, B. (2014). Using geographic information systems for radon exposure assessment in dwellings in the Oslo region, Norway. *Journal of Earth System Science*, *14*, 739–749. https://doi.org/10.5194/nhess-14-739-2014

Kreis, C., Lupatsch, J. E., Niggli, F., Egger, M., Kuehni, C. E., & Spycher, B. D. (2017). Space-Time Clustering of Childhood Leukemia: Evidence of an Association with ETV6-RUNX1 (TEL-AML1) Fusion. *Plos One*, *12*(1), e0170020–e0170020. https://doi.org/10.1371/journal.pone.0170020

Krestinina, L. Y., Davis, F. G., Schonfeld, S., Preston, D. L., Degteva, M., Epifanova, S., & Akleyev, A. V. (2013). Leukaemia incidence in the Techa River Cohort: 1953–2007. *British Journal of Cancer*, *109*(11), 2886–2893. https://doi.org/10.1038/bjc.2013.614

Kreuzer, M., Sobotzki, C., Fenske, N., Marsh, J. W., & Schnelzer, M. (2017). Leukaemia mortality and low-dose ionising radiation in the WISMUT uranium miner cohort (1946-2013). *Occupational and Environmental Medicine*, *74*(4), 252–258. https://doi.org/10.1136/oemed-2016-103795

Krewski, D., Lubin, J. H., Zielinski, J. M., Alavanja, M., Catalan, V. S., William Field, R., … Wilcox, H. B. (2006). A Combined Analysis of North American Case-Control Studies of Residential Radon and Lung Cancer. *Journal of Toxicology and Environmental Health, Part A*, *69*(7–8), 533–597. https://doi.org/10.1080/15287390500260945

Krille, L., Dreger, S., Schindel, R., Albrecht, T., Asmussen, M., Barkhausen, J., … Blettner, M. (2015). Risk of cancer incidence before the age of 15 years after exposure to ionising radiation from computed tomography: results from a German cohort study. *Radiation and Environmental Biophysics*, *54*(1), 1–12. https://doi.org/10.1007/s00411-014-0580-3

Kroll, M. E., Stiller, C. A., Murphy, M. F. G., & Carpenter, L. M. (2011). Childhood leukaemia and socioeconomic status in England and Wales 1976–2005: evidence of higher incidence in relatively affluent communities persists over time. *British Journal of Cancer*, *105*(11), 1783–1787. https://doi.org/10.1038/bjc.2011.415

Lachin, J. M. (2008). *Sample Size Evaluation for a Multiply Matched Case-Control Study Using the Score Test From a Conditional Logistic (Discrete Cox PH) Regression Model.* (27), 2509–2523. https://doi.org/10.1002/sim.3057

Laurier, D., Grosche, B., & Hall, P. (2002). Risk of Childhood Leukaemia in the Vicinity of Nuclear Installations: Findings and Recent Controversies. *Acta Oncologica*, *41*(1), 14–24. https://doi.org/10.1080/028418602317314019

Law, G. R., Kane, E. V., Roman, E., Smith, A., & Cartwright, R. (2000). Residential radon exposure and adult acute leukaemia. *The Lancet*, *355*(9218), 1888. https://doi.org/10.1016/S0140-6736(00)02300-X

Lee, C., Kim, K. P., Bolch, W. E., Moroz, B. E., & Folio, L. (2015). NCICT: a computational solution to estimate organ doses for pediatric and adult patients undergoing CT scans. *J. Radiol. Prot*, *35*. Retrieved from http://iopscience.iop.org/0952-4746/35/4/891

Leinonen, M. K., Miettinen, J., Heikkinen, S., Pitkäniemi, J., & Malila, N. (2017). Quality measures of the population-based Finnish Cancer Registry indicate sound data quality for solid malignant tumours. *European Journal of Cancer*, *77*, 31–39. https://doi.org/10.1016/j.ejca.2017.02.017

Leinonen, M. K., Rantanen, M., Pitkäniemi, J., & Malila, N. (2016). Coverage and accuracy of myeloproliferative and myelodysplastic neoplasms in the Finnish Cancer Registry. *Acta Oncologica*, *55*(6), 782–786. https://doi.org/10.3109/0284186X.2015.1127416

Leuraud, K., Richardson, D. B., Cardis, E., Daniels, R. D., Gillies, M., O'Hagan, J. A., … Kesminiene, A. (2015). Ionising radiation and risk of death from leukaemia and lymphoma in radiation-monitored workers (INWORKS): an international cohort study. *The Lancet Haematology*, *2*(7), e276–e281. https://doi.org/10.1016/S2352-3026(15)00094-0

Levy, S. E., & Myers, R. M. (2016). Advancements in Next-Generation Sequencing. *Annual Review of Genomics and Human Genetics*, *17*(1), 95–115. https://doi.org/10.1146/annurev-genom-083115-022413

Little, Kwon, D., Doi, K., Simon, S. L., Preston, D. L., Doody, M. M., … Sigurdson, A. J. (2014). Association of Chromosome Translocation Rate with Low Dose Occupational Radiation Exposures in U.S. Radiologic Technologists. *Radiation Research*, *182*(1), 1–17. https://doi.org/10.1667/RR13413.1

Little, Wakeford, R., Borrego, D., French, B., Zablotska, L. B., Adams, M. J., … Berrington de González, A. (2018). Leukaemia and myeloid malignancy among people exposed to low doses (<100 mSv) of ionising radiation during childhood: a pooled analysis of nine historical cohort studies. *The Lancet Haematology*, *5*(8), e346–e358. https://doi.org/10.1016/S2352-3026(18)30092-9

Little, Wakeford, R., Lubin, J. H., & Kendall, G. M. (2010). The statistical power of epidemiological studies analyzing the relationship between exposure to ionizing radiation and cancer, with special reference to childhood leukemia and natural background radiation. *Radiation Research*, *174*(3), 387–402. https://doi.org/10.1667/RR2110.1

Lorenzo-González, M., Ruano-Ravina, A., Peón, J., Piñeiro, M., & Barros-Dios, J. M. (2017). Residential radon in Galicia: a cross-sectional study in a radon-prone area. *Journal of Radiological Protection*, *37*(3), 728–741. https://doi.org/10.1088/1361-6498/aa7922

Lubin, J. H., Linet, M. S., Boice, J. D., Buckley, J., Conrath, S. M., Hatch, E. E., … Robison, L. L. (1998). Case-control study of childhood acute lymphoblastic leukemia and residential radon exposure. *Journal of the National Cancer Institute*, *90*(4), 294–300.

Lundell, M., & Holm, L.-E. (1996). Mortality from Leukemia after Irradiation in Infancy for Skin Hemangioma. *Radiation Research*, *145*(5), 595. https://doi.org/10.2307/3579278

Ma, X., Buffler, P. A., Selvin, S., Matthay, K. K., Wiencke, J. K., Wiemels, J. L., & Reynolds, P. (2002). Daycare attendance and risk of childhood acute lymphoblastic leukaemia. *British Journal of Cancer*, *86*(9), 1419–1424. https://doi.org/10.1038/sj.bjc.6600274

Madanat-Harjuoja, L. M., Pokhrel, A., Kivivuori, S. M., & Saarinen-Pihkala, U. M. (2014). Childhood cancer survival in Finland (1953-2010): A nation-wide population-based study: Survival after childhood cancer in Finland. *International Journal of Cancer*, *135*(9), 2129–2134. https://doi.org/10.1002/ijc.28844

Maged, A. F., Mokhtar, G. M., El-Tobgui, M. M., Gabbr, A. A., Attia, N. I., & Shady, M. M. A. (2000). Domestic radon concentration and childhood cancer study in Cairo, Egypt. *Journal of Environmental Science and Health, Part C*, *18*(2), 153–170. https://doi.org/10.1080/10590500009373519

Maia, A. T., Tussiwand, R., Cazzaniga, G., Rebulla, P., Colman, S., Biondi, A., & Greaves, M. (2004). Identification of preleukemic precursors of hyperdiploid acute lymphoblastic leukemia in cord blood. *Genes, Chromosomes and Cancer*, *40*(1), 38–43. https://doi.org/10.1002/gcc.20010

Mäkeläinen, Arvela, H., & Voutilainen, A. (2001). Correlations between radon concentration and indoor gamma dose rate, soil permeability and dwelling substructure and ventilation. *Science of the Total Environment*, *272*(1–3), 283–289. https://doi.org/10.1016/S0048-9697(01)00705-7

Mäkeläinen, Kinnunen, T., Reisbacka, H., Valmari, T., Arvela, H., & STUK. (2009). *Radon in Finnish dwellings - Sample Survey 2006 [in Finnish, abstract in English]*. Helsinki: Radiation and Nuclear Safety Authority.

Mäkeläinen, Moisio, S., Reisbacka, H., & Turtiainen, T. (2005). Indoor occupancy and radon exposure in Finland. In *Radioactivity in the Environment* (Vol. 7, pp. 687–693). https://doi.org/10.1016/S1569-4860(04)07084-6

Mäkeläinen, Valmari, Reisbacka, Kinnunen, & Arvela. (2010). Indoor radon and construction practices of Finnish homes from 20th to 21st century (Full paper for oral presentation S03-11). *Proceedings of Third European IRPA Congress 2010 June 14-16. Helsinki, Finland.*

Mansournia, M. A., Jewell, N. P., & Greenland, S. (2018). Case–control matching: effects, misconceptions, and recommendations. *European Journal of Epidemiology*, *33*(1), 5–14. https://doi.org/10.1007/s10654-017-0325-0

Marcotte, E. L., Thomopoulos, T. P., Infante-Rivard, C., Clavel, J., Petridou, E. T., Schüz, J., … Spector, L. G. (2016). Caesarean delivery and risk of childhood leukaemia: a pooled analysis from the Childhood Leukemia International Consortium (CLIC). *The Lancet Haematology*, *3*(4), e176–e185. https://doi.org/10.1016/S2352-3026(16)00002-8

Mathews, J. D., Forsythe, A. V., Brady, Z., Butler, M. W., Goergen, S. K., Byrnes, G. B., … Darby, S. C. (2013). Cancer risk in 680,000 people exposed to computed tomography scans in childhood or adolescence: data linkage study of 11 million Australians. *BMJ (Clinical Research Ed.)*, *346*(may21_1), f2360–f2360. https://doi.org/10.1136/bmj.f2360

McNally, R. J. Q., Alexander, F. E., & Bithell, J. F. (2006). Space-time clustering of childhood cancer in Great Britain: A national study, 1969–1993. *International Journal of Cancer*, *118*(11), 2840–2846. https://doi.org/10.1002/ijc.21726

Medical Birth Register - THL. (2016). Retrieved May 10, 2019, from https://thl.fi/en/web/thlfi-en/statistics/information-on-statistics/register-descriptions/newborns

Meinert, R., Kaletsch, U., Kaatsch, P., Schüz, J., & Michaelis, J. (1999). Associations between childhood cancer and ionizing radiation: results of a population-based case-control study in Germany. *Cancer Epidemiology Biomarkers & Prevention*, *8*(9), 793–799.

Metayer, C., Milne, E., Clavel, J., Infante-Rivard, C., Petridou, E., Taylor, M., … Buffler, P. A. (2013). The Childhood Leukemia International Consortium. *Cancer Epidemiology*, *37*(3), 336–347. https://doi.org/10.1016/j.canep.2012.12.011

Metayer, C., Milne, E., Dockerty, J. D., Clavel, J., Pombo-de-Oliveira, M. S., Wesseling, C., … Infante-Rivard, C. (2014). Maternal Supplementation with Folic Acid and Other Vitamins and Risk of Leukemia in Offspring: A Childhood Leukemia International Consortium Study. *Epidemiology*, *25*(6), 811–822. https://doi.org/10.1097/EDE.0000000000000141

Metayer, C., Petridou, E., Aranguré, J. M. M., Roman, E., Schüz, J., Magnani, C., … Clavel, J. (2016). Parental Tobacco Smoking and Acute Myeloid Leukemia: The Childhood Leukemia International Consortium. *American Journal of Epidemiology*, *184*(4), 261–273. https://doi.org/10.1093/aje/kww018

Meulepas, J. M., Ronckers, C. M., Smets, A. M. J. B., Nievelstein, R. A. J., Gradowska, P., Lee, C., … Hauptmann, M. (2018). Radiation Exposure from Pediatric CT Scans and Subsequent Cancer Risk in the Netherlands. *JNCI: Journal of the National Cancer Institute*. https://doi.org/10.1093/jnci/djy104

Miglioretti, D. L., Johnson, E., Williams, A., Greenlee, R. T., Weinmann, S., Solberg, L. I., … Smith-Bindman, R. (2013). The use of computed tomography in pediatrics and the associated radiation exposure and estimated cancer risk. *JAMA Pediatr*, *167*(8), 700–707. https://doi.org/10.1001/jamapediatrics.2013.311

Miles, J. C. H., Appleton, J. D., Rees, D. M., Green, B. M. R., Adlam, K. A. M., & Myers, A. H. (2007). *Indicative Atlas of Radon in England and Wales*. 36.

Morra, M. E., Kien, N. D., Elmaraezy, A., Abdelaziz, O. A. M., Elsayed, A. L., Halhouli, O., … Huy, N. T. (2017). Early vaccination protects against childhood leukemia: A systematic review and meta-analysis. *Scientific Reports*, *7*(1). https://doi.org/10.1038/s41598-017-16067-0

Mueller, W., & Gilham, C. (2015). Childhood leukemia and proximity to nuclear power plants: A systematic review and meta-analysis. *Journal of Cancer Policy*, *6*, 44–56. https://doi.org/10.1016/j.jcpo.2015.10.003

Muikku, Bly, R., Lahtinen, J., Lehtinen, M., Siiskonen, T., Turtiainen, T., … Vesterbacka, K. (2014). *Suomalaisten keskimääräinen efektiivinen annos Suomalaisten keskimääräinen efektiivinen annos*.

Muikku, M., Bly, R., Kurttio, P., Lahtinen, J., Lehtinen, M., Siiskonen, T., … Vesterbacka, K. (2014). *Suomalaisten keskimääräinen efektiivinen annos: Annoskakku 2012*. Säteilyturvakeskus.

Hunger SP, Mullighan CG. Acute Lymphoblastic Leukemia in Children. *New England Journal of Medicine*. 2015 Oct 15;373(16):1541-52. doi: 10.1056/NEJMra1400972

Nair, R. R. K., Rajan, B., Akiba, S., Jayalekshmi, P., Nair, M. K., Gangadharan, P., … Sugahara, T. (2009). Background radiation and cancer incidence in Kerala, India-Karanagappally cohort study. *Health Physics*, *96*(1), 55–66. https://doi.org/10.1097/01.HP.0000327646.54923.11

National Research Council. (1999). *Health Effects of Exposure to Radon: BEIR VI*. https://doi.org/10.17226/5499

Nguyen, K., Devidas, M., Cheng, S.-C., La, M., Raetz, E. A., Carroll, W. L., … Loh, M. L. (2008). Factors influencing survival after relapse from acute lymphoblastic leukemia: a Children's Oncology Group study. *Leukemia*, *22*(12), 2142–2150. https://doi.org/10.1038/leu.2008.251

Noshchenko, A. G., Bondar, O. Y., & Drozdova, V. D. (2010). Radiation-induced leukemia among children aged 0-5 years at the time of the Chernobyl accident. *International Journal of Cancer*. https://doi.org/10.1002/ijc.24834

Oksuzyan, S., Crespi, C. M., Cockburn, M., Mezei, G., Vergara, X., & Kheifets, L. (2015). Race/ethnicity and the risk of childhood leukaemia: a case–control study in California. *Journal of Epidemiology and Community Health*, *69*(8), 795–802. https://doi.org/10.1136/jech-2014-204975

Oren, O., Kebebew, E., & Ioannidis, J. P. A. (2019). Curbing Unnecessary and Wasted Diagnostic Imaging. *JAMA*, *321*(3), 245. https://doi.org/10.1001/jama.2018.20295

Paulsson, K., & Johansson, B. (2009). High hyperdiploid childhood acute lymphoblastic leukemia. *Genes, Chromosomes and Cancer*, *48*(8), 637–660. https://doi.org/10.1002/gcc.20671

Pearce, M. S., Salotti, J. A., Little, M. P., McHugh, K., Lee, C., Kim, K. P., … Berrington-De Gonzalez, A. B. (2012). Radiation exposure from CT scans in childhood and subsequent risk of leukaemia and brain tumours: A retrospective cohort study. *The Lancet*, *380*(9840), 499–505. https://doi.org/10.1016/S0140-6736(12)60815-0

Peres-Neto, P. R., Legendre, P., Dray, S., & Borcard, D. (2006). Variation partitioning of species data matrices: Estimation and comparison of fractions. *Ecology*, *87*(10), 2614–2625. https://doi.org/10.1890/0012-9658(2006)87[2614:VPOSDM]2.0.CO;2

Petridou, E. Th., Georgakis, M. K., Erdmann, F., Ma, X., Heck, J. E., Auvinen, A., … Skalkidou, A. (2018). Advanced parental age as risk factor for childhood acute lymphoblastic leukemia: results from studies of the Childhood Leukemia International Consortium. *European Journal of Epidemiology*, *33*(10), 965–976. https://doi.org/10.1007/s10654-018-0402-z

Piller, G. J. (2001). Leukaemia - a brief historical review from ancient times to 1950. *British Journal of Haematology*, *112*(2), 282–292. https://doi.org/10.1046/j.1365-2141.2001.02411.x

Pritchard-Jones, K. (1996). Genetics of childhood cancer. *Cancer in Children*, 20.

Pui, C.-H., & Relling, M. V. (2000). Topoisomerase II inhibitor-related acute myeloid leukaemia. *British Journal of Haematology*, *109*(1), 13–23. https://doi.org/10.1046/j.1365-2141.2000.01843.x

Raaschou-Nielsen, O., Andersen, C. E., Andersen, H. P., Gravesen, P., Lind, M., Schüz, J., & Ulbak, K. (2008). Domestic radon and childhood cancer in Denmark. *Epidemiology (Cambridge, Mass.)*, *19*(4), 536–543. https://doi.org/10.1097/EDE.0b013e318176bfcd

Rajaraman, P., Simpson, J., Neta, G., Berrington de Gonzalez, A., Ansell, P., Linet, M. S., … Roman, E. (2011). Early life exposure to diagnostic radiation and ultrasound scans and risk of childhood cancer: case-control study. *BMJ*, *342*(feb10 1), d472–d472. https://doi.org/10.1136/bmj.d472

Register of Congenital malformations - THL. (2015). Retrieved May 10, 2019, from https://thl.fi/en/web/thlfi-en/statistics/information-on-statistics/register-descriptions/register-of-congenital-malformations

Reisz, J. A., Bansal, N., Qian, J., Zhao, W., & Furdui, C. M. (2014). Effects of Ionizing Radiation on Biological Molecules—Mechanisms of Damage and Emerging Methods of Detection. *Antioxidants & Redox Signaling*, *21*(2), 260–292. https://doi.org/10.1089/ars.2013.5489

Rericha, V., Kulich, M., Rericha, R., Shore, D. L., & Sandler, D. P. (2006). Incidence of leukemia, lymphoma, and multiple myeloma in Czech uranium miners: a case-cohort study. *Environmental Health Perspectives*, *114*(6), 818–822. https://doi.org/10.1289/ehp.8476

Revzan, K. L., & Fisk, W. J. (1992). Modeling Radon Entry into Houses with Basements: The Influence of Structural Factors. *Indoor Air*, *2*(1), 40–48. https://doi.org/10.1111/j.1600-0668.1992.05-21.x

Richardson, R. B., Eatough, J. P., & Henshaw, D. L. (1991). Dose to red bone marrow from natural radon and thoron exposure. *The British Journal of Radiology*, *64*(763), 608–624. https://doi.org/10.1259/0007-1285-64-763-608

Ron, E., Modan, B., & Boice, J. D. (1988). Mortality after radiotherapy for ringworm of the scalp (abstract only). *American Journal of Epidemiology*, *127*(4), 713–725.

Rosenberg, S. A., & Kaplan, H. S. (1985). The evolution and summary results of the Stanford randomized clinical trials of the management of Hodgkin's disease: 1962–1984. *International Journal of Radiation Oncology*Biology*Physics*, *11*(1), 5–22. https://doi.org/10.1016/0360-3016(85)90357-8

Ross, J. A., Severson, R. K., Swensen, A. R., Pollock, B. H., Gurney, J. G., & Robison, L. L. (1999). Seasonal variations in the diagnosis of childhood cancer in the United States. *British Journal of Cancer*, *81*(3), 549–553. https://doi.org/10.1038/sj.bjc.6690729

Ross, Potter, J. D., Shu, X.-O., Reaman, G. H., Lampkin, B., & Robison, L. L. (1997). Evaluating the relationships among maternal reproductive history, birth characteristics, and infant leukemia: A report from the children's cancer group. *Annals of Epidemiology*, *7*(3), 172–179. https://doi.org/10.1016/S1047-2797(97)00012-4

Rothman, K. J. (2014). Six Persistent Research Misconceptions. *Journal of General Internal Medicine*, *29*(7), 1060–1064. https://doi.org/10.1007/s11606-013-2755-z

Rubnitz, J. E. (2017). Current Management of Childhood Acute Myeloid Leukemia. *Pediatric Drugs*, *19*(1), 1–10. https://doi.org/10.1007/s40272-016-0200-6

Safavi, S., Forestier, E., Golovleva, I., Barbany, G., Nord, K. H., Moorman, A. V., … Paulsson, K. (2013). Loss of chromosomes is the primary event in near-haploid and low-hypodiploid acute lymphoblastic leukemia. *Leukemia*, *27*(1), 248–250. https://doi.org/10.1038/leu.2012.227

Schäfer, D., Olsen, M., Lähnemann, D., Stanulla, M., Slany, R., Schmiegelow, K., … Fischer, U. (2018). Five percent of healthy newborns have an *ETV6-RUNX1* fusion as revealed by DNA-based GIPFEL screening. *Blood*, *131*(7), 821–826. https://doi.org/10.1182/blood-2017-09-808402

Schipler, A., Mladenova, V., Soni, A., Nikolov, V., Saha, J., Mladenov, E., & Iliakis, G. (2016). Chromosome thripsis by DNA double strand break clusters causes enhanced cell lethality, chromosomal translocations and 53BP1-recruitment. *Nucleic Acids Research*, *44*(16), 7673–7690. https://doi.org/10.1093/nar/gkw487

Schmidhuber, J. (2015). Deep Learning in Neural Networks: An Overview. *Neural Networks*, *61*, 85–117. https://doi.org/10.1016/j.neunet.2014.09.003

Schmiegelow K, Lausten Thomsen U, Baruchel A, Pacheco CE, Pieters R, Pombo-de-Oliveira MS, Andersen EW, Rostgaard K, Hjalgrim H, Pui CH. High concordance of subtypes of childhood acute lymphoblastic leukemia within families: lessons from

sibships with multiple cases of leukemia. *Leukemia*. 2012 Apr;26(4):675-81. doi: 10.1038/leu.2011.274

Schuz, J., & Ahlbom, A. (2008). Exposure to electromagnetic fields and the risk of childhood leukaemia: a review. *Radiation Protection Dosimetry*, *132*(2), 202–211. https://doi.org/10.1093/rpd/ncn270

Shibamoto, Y., & Nakamura, H. (2018). Overview of Biological, Epidemiological, and Clinical Evidence of Radiation Hormesis. *International Journal of Molecular Sciences*, *19*(8), 2387. https://doi.org/10.3390/ijms19082387

Shore, Beck, H. L., Boice, J. D., Caffrey, E. A., Davis, S., Grogan, H. A., … Dauer, L. T. (2018). Implications of recent epidemiologic studies for the linear nonthreshold model and radiation protection. *Journal of Radiological Protection*, *38*(3), 1217–1233. https://doi.org/10.1088/1361-6498/aad348

Shore, Moseson, M., Harley, N., & Pasternack, B. S. (2003). Tumors and other diseases following childhood X-ray treatment for ringworm of the scalp (tinea capitis). *Health Physics*, *85*(4), 404–408. https://doi.org/10.1097/00004032-200310000-00003

Shu, X. O., Potter, J. D., Linet, M. S., Severson, R. K., Han, D., Kersey, J. H., … Robison, L. L. (2002). Diagnostic X-rays and ultrasound exposure and risk of childhood acute lymphoblastic leukemia by immunophenotype. *Cancer Epidemiol Biomarkers Prev*, *11*(2), 177–185.

Sigurdson, A. J., Ha, M., Hauptmann, M., Bhatti, P., Sram, R. J., Beskid, O., … Tucker, J. D. (2008). International study of factors affecting human chromosome translocations. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, *652*(2), 112–121. https://doi.org/10.1016/j.mrgentox.2008.01.005

Spearman, C. (1904). The Proof and Measurement of Association between Two Things. *The American Journal of Psychology*, 15(1), 72-101. doi:10.2307/1412159

Spector, L. G. (2005). Maternal Diet and Infant Leukemia: The DNA Topoisomerase II Inhibitor Hypothesis: A Report from the Children's Oncology Group. *Cancer Epidemiology Biomarkers & Prevention*, *14*(3), 651–655. https://doi.org/10.1158/1055-9965.EPI-04-0602

Spix, C., Grosche, B., Bleher, M., Kaatsch, P., Scholz-Kreisel, P., & Blettner, M. (2017). Background gamma radiation and childhood cancer in Germany: an ecological study. *Radiation and Environmental Biophysics*, *56*(2), 127–138. https://doi.org/10.1007/s00411-017-0689-2

Spycher, B. D., Lupatsch, J. E., Zwahlen, M., Röösli, M., Niggli, F., Grotzer, M. A., … the Swiss National Cohort Study Group. (2015). Background Ionizing Radiation and the Risk of Childhood Cancer: A Census-Based Nationwide Cohort Study. *Environmental Health Perspectives*, *123*(6), 622–628. https://doi.org/10.1289/ehp.1408548

Steinbuch, M., Weinberg, C. R., Buckley, J. D., Robison, L. L., & Sandler, D. P. (1999). Indoor residential radon exposure and risk of childhood acute myeloid leukaemia. *British Journal of Cancer*, *81*(5), 900–906. https://doi.org/10.1038/sj.bjc.6690784

Steliarova-Foucher, E., Colombet, M., Ries, L. A. G., Moreno, F., Dolya, A., Bray, F., … Steliarova-Foucher, E. (2017). International incidence of childhood cancer, 2001–10: a population-based registry study. *The Lancet Oncology*, *18*(6), 719–731. https://doi.org/10.1016/S1470-2045(17)30186-9

Stevens, W., Thomas, D. C., Lyon, J. L., Till, J. E., Kerber, R. A., Simon, S. L., … Preston-Martin, S. (1990). Leukemia in Utah and radioactive fallout from the Nevada test site. A case-control study. *JAMA*, *264*(5), 585–591.

Stieglitz, E., & Loh, M. L. (2013). Genetic predispositions to childhood leukemia. *Therapeutic Advances in Hematology*, *4*(4), 270–290.

Stovall M, Weathers R, Kasper C, Smith SA, Travis L, Ron E, Kleinerman R. Dose reconstruction for therapeutic and diagnostic radiation exposures: use in epidemiological studies. *Radiat Res.* 2006 Jul;166(1 Pt 2):141-57.

Suutari, J. (2015). *The numbers of radiological studies and procedures in 2015 [in Finnish, abstract in English]*. STUK.

Tao, Z., Akiba, S., Zha, Y., Sun, Q., Zou, J., Li, J., … Wei, L. (2012). Cancer and non-cancer mortality among Inhabitants in the High Background Radiation Area of Yangjiang, China (1979–1998): *Health Physics*, *102*(2), 173–181. https://doi.org/10.1097/HP.0b013e31822c7f1e

Kaohsiung Leukemia Research Group, Liu, C., Hsu, Y.-H., Wu, M.-T., Pan, P.-C., Ho, C.-K., … Christiani, D. C. (2009). Cured meat, vegetables, and bean-curd foods in relation to childhood acute leukemia risk: A population based case-control study. *BMC Cancer*, *9*(1). https://doi.org/10.1186/1471-2407-9-15

Tilastokeskus - Luokitukset - Classification of Socio-economic Groups 1989. (2019).

Toft, N., Birgens, H., Abrahamsson, J., Griškevičius, L., Hallböök, H., Heyman, M., … Schmiegelow, K. (2018). Results of NOPHO ALL2008 treatment for patients aged 1–45 years with acute lymphoblastic leukemia. *Leukemia*, *32*(3), 606–615. https://doi.org/10.1038/leu.2017.265

Tong, J., Qin, L., Cao, Y., Li, J., Zhang, J., Nie, J., & An, Y. (2012). Environmental radon exposure and childhood leukemia. *Journal of Toxicology and Environmental Health. Part B, Critical Reviews*, *15*(5), 332–347. https://doi.org/10.1080/10937404.2012.689555

Tucker, M. A., Meadows, A. T., Boice, J. D., Stovall, M., Oberlin, O., Stone, B. J., … Fraumeni, J. F. (1987). Leukemia after therapy with alkylating agents for childhood cancer (abstract). *Journal of the National Cancer Institute*, *78*(3), 459–464.

Turner, M. C., Wigle, D. T., & Krewski, D. (2010). Residential Pesticides and Childhood Leukemia: A Systematic Review and Meta-Analysis. *Environmental Health Perspectives*, *118*(1), 33–41. https://doi.org/10.1289/ehp.0900966

UK Childhood Cancer Study Investigators. (2002). The United Kingdom Childhood Cancer Study of exposure to domestic sources of ionising radiation: 1: radon gas. *British Journal of Cancer*, *86*(11), 1721–1726. https://doi.org/10.1038/sj.bjc.6600276

UK Childhood Cancer Study Investigators. (2002). The United Kingdom Childhood Cancer Study of exposure to domestic sources of ionising radiation: 2: gamma radiation. *British Journal of Cancer*, *86*(11), 1727–1731. https://doi.org/10.1038/sj.bjc.6600277

United Nations Scientific Committee on the Effects of Atomic Radiation (Ed.). (2000). *Sources and effects of ionizing radiation: United Nations Scientific Committee on the Effects of Atomic Radiation: UNSCEAR 2000 report to the General Assembly, with scientific annexes*. New York: United Nations.

United Nations Scientific Committee on the Effects of Atomic Radiation. (2010). *Sources and effects of ionizing radiation: United Nations Scientific Committee on the Effects of Atomic Radiation: UNSCEAR 2008 report*. New York: United Nations.

United Nations. Scientific Committee on the Effects of Atomic Radiation. (2013). *UNSCEAR 2013 Report: Effects of radiation exposure on children*. New York.

Valmari, T., Arvela, H., & Reisbacka, H. (2012). Radon in Finnish apartment buildings. *Radiation Protection Dosimetry*, *152*(1–3), 146–149. https://doi.org/10.1093/rpd/ncs211

Wakeford, R., Kendall, G. M., & Little, M. P. (2009). The proportion of childhood leukaemia incidence in Great Britain that may be caused by natural background ionizing radiation. *Leukemia*, *23*(4), 770–776. https://doi.org/10.1038/leu.2008.342

Wall, B. F., & Hart, D. (1997). Revised radiation doses for typical X-ray examinations. *T He British Journal of Radiology*, 3.

Wallace, A. D., Francis, S. S., Ma, X., McKean-Cowdin, R., Selvin, S., Whitehead, T. P., … Metayer, C. (2018). Allergies and Childhood Acute Lymphoblastic Leukemia: A Case–Control Study and Meta-analysis. *Cancer Epidemiology Biomarkers & Prevention*, *27*(10), 1142–1150. https://doi.org/10.1158/1055-9965.EPI-17-0584

Weinberg, C. R., Moledor, E. S., Umbach, D. M., & Sandler, D. P. (1996). Imputation for Exposure Histories with Gaps, under an Excess Relative Risk Model. *Epidemiology*, *7*(5), 490–497.

Westerbeek, R., Blair, V., Eden, O., Kelsey, A., Stevens, R., Will, A., … Birch, J. (1998). Seasonal variations in the onset of childhood leukaemia and lymphoma. *British Journal of Cancer*, *78*(1), 119–124. https://doi.org/10.1038/bjc.1998.452

Wigle, D. T., Turner, M. C., & Krewski, D. (2009). A Systematic Review and Meta-analysis of Childhood Leukemia and Parental Occupational Pesticide Exposure. *Environmental Health Perspectives*, *117*(10), 1505–1513. https://doi.org/10.1289/ehp.0900582

World Health Organization. (2009). *WHO handbook on indoor radon: a public health perspective*. Geneva, Switzerland: World Health Organization.

World Health Organization. (2016). *International statistical classification of diseases and related health problems: 10th revision ICD-10*. Geneva: World Health Organization.

Yeazel, M. W., Buckley, J. D., Woods, W. G., Ruccione, K., & Robison, L. L. (1995). History of maternal fetal loss and increased risk of childhood acute leukemia at an early age. A report from the Childrens Cancer Group. *Cancer*, *75*(7), 1718–1727.

Yoshinaga, S., Tokonami, S., Akiba, S., Nitta, H., & Kabuto, M. (2005). Case-control study of residential radon and childhood leukemia in Japan: results from preliminary analyses. *International Congress Series*, *1276*, 233–235. https://doi.org/10.1016/j.ics.2004.09.050

Zierhut, H., Linet, M. S., Robison, L. L., Severson, R. K., & Spector, L. (2012). Family history of cancer and non-malignant diseases and risk of childhood acute lymphoblastic leukemia: A Children's Oncology Group Study. *Cancer Epidemiology*, *36*(1), 45–51. https://doi.org/10.1016/j.canep.2011.06.004

PUBLICATIONS

PUBLICATION I

Background radiation and childhood leukemia: A nationwide register-based case-control study

Atte Nikkilä, Sini Erme, Hannu Arvela, Olli Holmgren, Jani Raitanen, Olli Lohi, Anssi Auvinen

International Journal of Cancer 139(9), 1975-1982
https://doi.org/10.1002/ijc.30264

IJC
International Journal of Cancer

Background radiation and childhood leukemia: A nationwide register-based case-control study

Atte Nikkilä[1], Sini Erme[2], Hannu Arvela[3], Olli Holmgren[3], Jani Raitanen[4,5], Olli Lohi[2] and Anssi Auvinen[2,3,4]

[1] School of Medicine, University of Tampere, Tampere, Finland
[2] Tampere Center for Child Health Research, Tampere University Hospital and University of Tampere, Tampere, Finland
[3] STUK—Radiation and Nuclear Safety Authority, Environmental Surveillance, Helsinki, Finland
[4] School of Health Sciences, University of Tampere, Tampere, Finland
[5] UKK Institute for Health Promotion, Tampere, Finland

High doses of ionizing radiation are an established cause of childhood leukemia. However, substantial uncertainty remains about the effect of low doses of radiation, including background radiation and potential differences between genetic subgroups of leukemia have rarely been explored. We investigated the effect of the background gamma radiation on childhood leukemia using a nationwide register-based case-control study. For each of the 1,093 cases, three age- and gender matched controls were selected ($N = 3,279$). Conditional logistic regression analyses were adjusted for confounding by Down syndrome, birth weight (large for gestational age), and maternal smoking. Complete residential histories and previously collected survey data of the background gamma radiation in Finland were used to assess the exposure of the study subjects to indoor and outdoor gamma radiation. Overall, background gamma radiation showed a non-significant association with the OR of childhood leukemia (OR 1.01, 95% CI 0.97, 1.05 for 10 nSv/h increase in average equivalent dose rate to red bone marrow). In subgroup analyses, age group 2–<7 years displayed a larger effect (OR 1.27, 95% CI 1.01, 1.60 for 1 mSv increase in equivalent cumulative dose to red bone marrow). Suggestive difference in OR by genetic subtype was found. Our results provide further support to the notion that low doses of ionizing radiation increase the risk for childhood leukemia, particularly at age 2–<7 years. Our findings suggest a larger effect of radiation on leukemia with high hyperpdiploidy than other subgroups, but this result requires further confirmation.

Childhood leukemia is a heterogeneous disease consisting of distinct clinical subtypes and characterized by specific chromosomal translocations or mutations.[1] Acute lymphoblastic leukemia (ALL), which encompasses around 85% of all childhood leukemias, has two major genetic subtypes both comprising ~25–30% of the cases: high hyperdiploid (HeH) and TEL-AML1 (TA) fusion precursor B-ALL.[1]

The etiology and environmental risk factors for childhood leukemia remain largely unknown. The few established risk factors include high doses of ionizing radiation, Down syndrome, and large birth weight for ALL.[2–5]

Based on studies of atomic bomb survivors and other populations, radiation is known to cause leukemia.[2] In general, children have a larger risk per dose unit than adults.[2] While the leukemogenic effect of moderate-to-high-dose radiation is well established, low doses of background radiation have also been suggested to increase leukemia risk though the shape of the dose-response remains uncertain implying a small incremental risk from the natural background gamma radiation.[6,7]

We set out to investigate the relationship between background gamma radiation and risk of childhood leukemia based on a nationwide register-based case-control study. We hypothesized that low-dose radiation from background gamma radiation will pose a risk of childhood leukemia in accordance with the linear no-threshold model.[8] We also evaluated potential differences in risk (effect modification) by age group and the genetic subtypes of ALL. We estimated lifetime cumulative doses and average dose rates to the red bone marrow (RBM) from background radiation using information on indoor and outdoor survey data and housing types during the lifespans of the children. Different subtypes of leukemia in terms of cell type, genetic aberrations and age at diagnosis were separately analyzed.

Key words: children, leukemia, background radiation, cancer epidemiology
Additional Supporting Information may be found in the online version of this article.
Grant sponsors: Tampereen Lastenklinikan Tukiry; Tellervo Edessalo Foundation
DOI: 10.1002/ijc.30264
History: Received 8 Feb 2016; Accepted 15 June 2016; Online 12 July 2016
Correspondence to: Atte Nikkilä, Finn-Medi 3, Biokatu 10, 33520 Tampere, Finland, Tel: +358-50-318-6254, E-mail: atte.nikkila@uta.fi

Methods

Study design

We used a matched, register-based case-control study design. Cases included all children diagnosed with leukemia

What's new?

While exposure to moderate and high levels of ionizing radiation is an established risk factor for childhood leukemia, whether low-level exposure also contributes to this risk remains uncertain. In this investigation of more than 1,000 Finnish patients with complete residential histories, no significant association was observed between background radiation and overall risk of childhood leukemia. Subgroup analyses, however, revealed a significant elevation in risk for individuals age 2–<7 years, the age group with the highest incidence of acute lymphocytic leukemia (ALL) in Finland. Of genetic subtypes of ALL, risk appeared to be highest for high hyperdiploidy.

(M9800–M9948 in ICD-O-3) in Finland in 1990–2011, identified from the Finnish Cancer Registry ($N = 1100$). Each case was matched by gender and year of birth to three controls at the Population Register Center. Each control was assigned a reference date identical to the diagnosis date of the respective case, so that the controls were of similar age as their case at the date up to which exposure was analyzed (end of the exposure period). A 2-year minimum latency period was assumed in the main analysis based on previous knowledge resulting in zero exposure for subjects younger than 2 years at the reference date.

Data collection

The data obtained from the Finnish Cancer Register contained the diagnosis (ICD-O-3 code), month and year of the diagnosis and date of birth. The diagnosis was classified into three categories: acute lymphoblastic leukemia (ALL), acute myeloid leukemia (AML), and others. Seven cases without a valid personal identification number or prohibiting the use of their information at the population register were excluded ($3 + 4 = 7$) (Fig. 1), leaving 1093 cases for the analysis. For the matched cases and controls, Population Register Center provided complete residential histories from birth to the reference date, including moving dates, municipalities of residence, as well as coordinates and building identification codes for each residence. Of the study subjects, 51.9% (52.1% of cases and 51.9% of controls) had lived in only one dwelling and 10.1% of study subjects (9.5% of cases and 10.3% of controls) in 4 or more dwellings (Supporting Information Fig. S1). Building codes were missing for 3.6% (4.4% for cases and 3.3% for controls) of the buildings and in such cases; the typical building type of the municipality was assigned (Supporting Information Fig. S2).

Data on diagnosis of Down syndrome were obtained from Register of Congenital Malformations by the National Institute of Health and Welfare. Data on gestational weeks and birth weights were obtained from the Medical Birth Register for all but 184 (18.6%) cases and 511 (15.6%) controls. Large for gestational age (LGA) was defined as birth weight exceeding the 90th percentile given the gestational weeks.[9] Maternal smoking data was obtained from the Medical Birth Register and data was missing for 200 cases (18.3%) and 563 controls (17.2%). The implications of these missing data can be visualized in Figure 1.

Genetic data for cases were obtained from the university hospitals in charge of treatment of all childhood leukemias in Finland (Helsinki, Tampere, Turku, Oulu, and Kuopio). The genetic aberrations were classified as TEL-AML1, high hyperdiploid (HeH), other abnormalities and normal (no detected abnormalities). Data were missing or unavailable for 146 (13.3%) patients, mostly due to limited availability of genetic tests during the early study period. The relatively low proportion of TEL-AML1 cases in our material (9.2%) resulted from the fact that testing for this genetic subtype began in Finland after year 1998. The proportion of TEL-AML1 cases remains relatively stable (14.1%) from the year 1998 onwards.

The data on dose rates of natural background gamma radiation outdoors was obtained from STUK—Radiation and Nuclear Safety Authority, encoded in a map of 8 km × 8 km squares based on a nationwide mobile survey carried out in 1978–1980[10] (Fig. 2). We also obtained dose rates calculated for each of the 355 Finnish municipalities from STUK. The nationwide average dose rate outdoors was 51 nSv/h. For the residencies abroad ($n = 63$, 0.8% of all residencies), we used world's average natural background radiation value reported by UNSCEAR (55.3 nSv h^{-1} effective dose rate).[8] The doses of gamma radiation indoors are based on measurements in 346 randomly chosen dwellings. The national average dose rate in houses was 41 nSv h^{-1} and in flats 70 nSv h^{-1}.[10] The difference is mainly due to the concrete used as the building material in blocks of flats increasing gamma radiation levels. The indoor rates correlate with local outdoor gamma radiation levels. These correlations were utilized for converting local outdoor dose rates to indoor dose rates.

The data on Chernobyl fallout was collected by a nationwide mobile survey in 1986–1987.[11] A square map of the Chernobyl fallout representing the Cs-137 activity in October 1, 1986 was transformed into dose rates (Supporting Information Fig. S3). Doses were calculated with a function of dose rate consisting of three individual parts. The first ranged 2 weeks from the initial fallout and, like the second part, was fitted strictly on measured data. Correction for radioactive decay of all relevant short-lived radionuclides was also applied. The third part continued from 18 months after the accident based on exponential decay representing the effective decay and washout of the fallout nuclides, mainly Cs-137. Results from fixed radiation monitoring stations were utilized. Concerning the doses from Chernobyl fallout

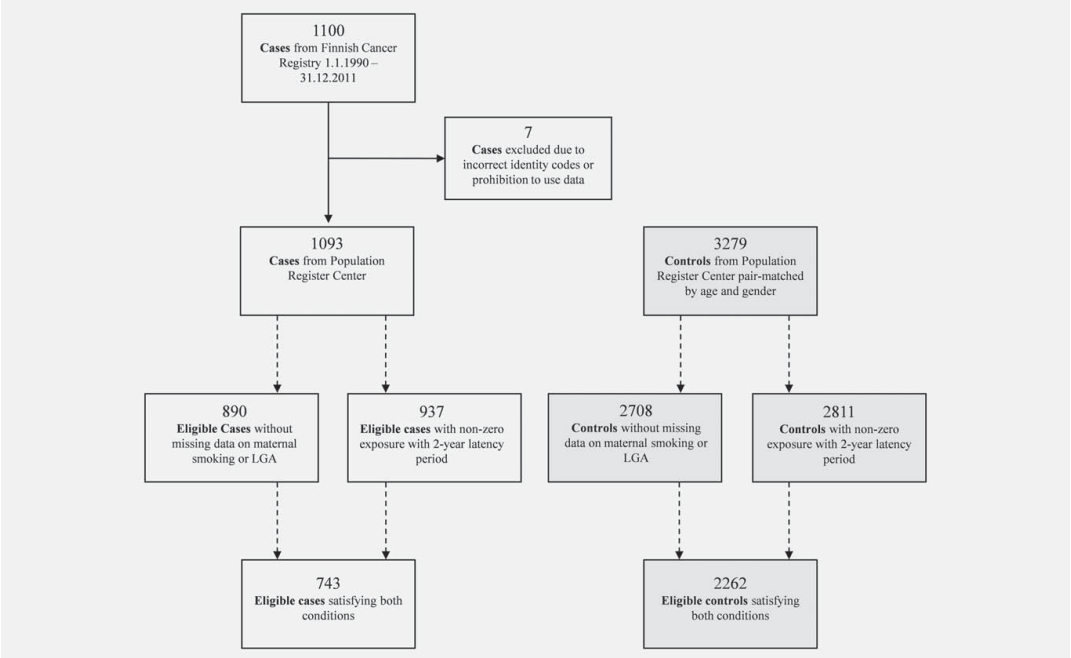


Figure 1. Flow chart depicting the choice of cases and controls and the necessary exclusions. The boxes at the bottom represent the scenario in which adjusting was possible.

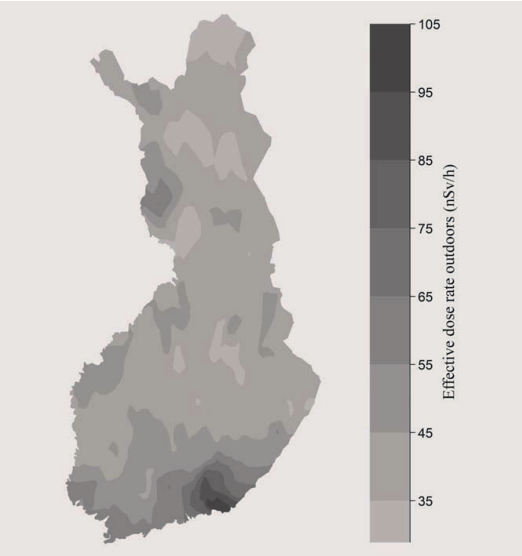


Figure 2. Natural background radiation outdoors in Finland. Color coded intervals depict effective dose rate outdoors and the values are in nSv/h.

received indoors, separate shielding factors for houses and blocks of flats were used. The shielding factor for blocks of flats (0.04) is markedly higher than that for houses (0.37).[10] Cosmic radiation was not included in our analysis, because Finland is a fairly flat country with little variation in dose rate. Thus, it would have introduced no variation between cases and controls. The altitudes of the residencies of study subjects were analyzed to confirm the assumption.

To account for the markedly different dose rates and shielding effect by dwelling type, we classified each residential building as a house (single family and terraced houses) or a block of flats. The Population Register Center provided the information for 95% (94% for cases and 95% for controls) of the dwellings, and the classification was based on available documentation on the type of the dwelling or on the number of floors and dwellings in the building, and the presence of an elevator. We obtained data on the types of dwellings of children from Statistics Finland to model the residencies, which could not be classified. With these data, we could define average dwelling type coefficient for children of different ages during different calendar years and therefore better approximate the conversion from outdoor absorbed dose rates to the indoor dose rates.

The percentage of time spent indoors (occupancy) was modeled according to a Finnish study providing age group

Table 1. Characteristics of cases and controls

	Cases ($n = 1,093$)	Controls ($n = 3,279$)	Total ($n = 4,372$)
Gender			
Male	568 (52.0%)	1,704 (52.0%)	2,272 (52.0%)
Female	525 (48.0%)	1,575 (48.0%)	2,100 (48.0%)
Large for gestational age			
Yes	121 (13.3%)	275 (9.9%)	396 (10.8%)
No	788 (86.7%)	2,493 (90.1%)	3,281 (89.2%)
Missing	184	511	695
Smoking during pregnancy			
Yes	151 (16.9%)	420 (15.5%)	571 (15.8%)
No	742 (83.1%)	2,296 (84.5%)	3,038 (84.2%)
Missing	200	563	763
Down syndrome			
Yes	40 (3.7%)	2 (0.1%)	42 (1.0%)
No	1,053 (96.3%)	3,277 (99.9%)	4,330 (99.0%)
Age at diagnosis, years			
0–<2	156 (14.3%)		
2–<7	605 (55.5%)		
7–<15	332 (33.4%)		
Leukemia type			
ALL	885 (81.1%)		
AML	142 (13.0%)		
Other	66 (5.9%)		
Genetic subtype			
TEL-AML1	87 (9.2%)		
HeH[1]	227 (24.0%)		
Other abnormalities	382 (40.3%)		
Normal	251 (26.5%)		
Missing	146		

[1]Cases with both HeH and TEL-AML1 properties were included in the HeH category.

specific estimates.[12] For foreign residencies, we used occupancy estimates by UNSCEAR (0.8).[8]

Data analysis

We used a coefficient of 0.7 Sv/Gy to convert absorbed dose rates from environmental measurements (from natural background radiation and Chernobyl fallout) to effective dose rates.[8] We then applied the occupancy coefficient and coefficient of dwelling type to obtain effective dose rates for each residency by area and age.[10] Cumulative effective doses for each residential period were calculated as time integrals of dose rates, and summed to obtain lifetime cumulative effective doses. The average effective dose rates for each person was calculated by dividing cumulative dose by total residential history up to a date 2 years prior to diagnosis or reference date.

We calculated equivalent RBM doses from effective doses in the same manner as Kendall et al. based on the results published by Saito et al and Petoussi et al.[13–15] We assumed a linear relationship with age by converting adult effective dose rates to children of different ages instead of assuming a cubic root function in relation to the weight of the growing child.

Statistical methods

Odds ratios and their confidence intervals were calculated using conditional logistic regression. We used 5% as the level of significance and all p values are two-sided. Crude ORs were used if adjustment did not alter the OR more than 0.05 U. Smoothed dose-response plots are plotted with Bézier splines and the point estimates were calculated with the floating absolute risk method.[16] Eight subgroup analyses were

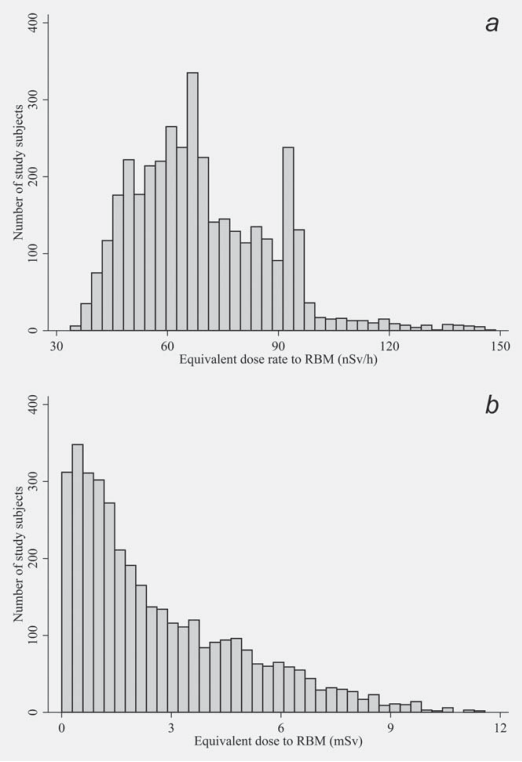


Figure 3. Average equivalent dose rates (*a*) and Cumulative equivalent doses (*b*) to RBM from natural background radiation and Chernobyl fallout.

performed and no correction for multiple testing was used. Effect modification *i.e.*, difference between subgroups was assessed based on the significance of an interaction term added to a model with the main effects. Statistical power calculations indicated that the material was sufficient for detecting a linear dose-response with OR of 1.06 or greater per 10 nSv h^{-1} increase in dose rate with statistical power of 80%.[17]

Ethical approval

The ethical committee of Pirkanmaa Hospital District reviewed the study protocol (tracking number R14074) and in accordance with the Finnish regulation, no informed consent was required for a register-based study. We obtained the permission to use data from Finnish Cancer Register and Medical Birth Register from the National Institute of Health and Welfare.

Results

In a nationwide register-based study, we identified a total of 1,093 leukemia cases diagnosed between 1990 and 2011 in Finland. The majority of the cases were acute lymphoblastic leukemias (81.1%), followed by acute myeloid leukemia (13.0%) (Table 1). The median age at diagnosis was 4.52 years (IQR 2.72, 8.23 years) and for study subjects with non-zero radiation exposure, 5.23 years (IQR 3.44, 9.05) (Supporting Information Fig. S4). Of the specific genetic alterations, high hyperdiploidy (HeH) and TEL-AML1 (TA) fusion gene were the most common abnormalities.

The median equivalent dose rate to red bone marrow (from natural background radiation and Chernobyl fallout combined) was 67.2 nSv h^{-1} for cases and 66.4 nSv h^{-1} for controls. The average equivalent dose rate to red bone marrow (Fig. 3) showed peaks at 67 and 94 nSv h^{-1}, due to the large number of houses and flats in the Helsinki metropolitan area (19.4% of the dwellings). The distribution of cumulative equivalent dose to RBM (Fig. 3) followed a log-normal distribution, with the smoother shape due to contribution from several past dwellings. The median cumulative equivalent dose to RBM for the cases was 1.96 mSv and for the controls 1.90 mSv. The median dose received only from Chernobyl fallout was 0.1 mSv for both the cases and the controls and the corresponding average dose rate was 2.0 nSv h^{-1}. The distributions of the dose rates from the Chernobyl fallout and natural background gamma both followed a log-normal distribution. The median altitude for the cases' residences was 82 m (IQR 23 m, 108 m) and for the controls 74 m (IQR 21 m, 104 m).

A very high leukemia OR was observed for children with Down syndrome (OR 60.0, 95% CI 14.5, 248) and an increased OR was also related to LGA (OR 1.44, 95% CI 1.14, 1.81). For maternal smoking during pregnancy, a non-significantly elevated OR was found (OR 1.15, 95% CI 0.94, 1.42).

The overall OR for leukemia was 1.01 (95% CI 0.97, 1.05) for each 10 nSv h^{-1} increase in the equivalent dose rate to RBM, and for a 1 mSv increase in cumulative equivalent dose to RBM, the OR was 0.97 (95% CI 0.89, 1.06) (Table 2). When divided into exposure quartiles, the main OR increase was between the lowest quarter (reference) and the two intermediate fourths (Fig. 4).

There was no statistically significant difference between the major cell types of leukemia in OR related to background radiation, though a higher point estimate was seen for ALL than AML. However, OR related to background radiation varied significantly by age at reference date (interaction $p = 0.005$). In the analysis by age group, a significantly elevated OR (OR 1.05 per 10 nSv h^{-1}, 95% CI 1.00, 1.10 for dose rate and OR 1.27, 95% CI 1.01, 1.60 per 1 mSv for cumulative dose) was found for ages 2–<7 years, the time of the peak in incidence of ALL and largest numbers of cases. The dose-response for the younger age group is shown in Figure 4.

For the major genetic subtypes of ALL, cases with a HeH genotype showed a significant increase in OR with dose rate from background radiation (OR 1.11 per 10 nSv h^{-1}, 95% CI 1.02, 1.21), though not with cumulative dose (OR 1.30, 95% CI 0.94, 1.80 per mSv). Yet, the risk estimates did not vary

Table 2. Odds ratios for the relationship between incremental increases in average dose rate and cumulative dose to red bone marrow and childhood leukemia

		Average equivalent dose rate—Increase of 10 nSv/h		Cumulative equivalent dose—Increase of 1 mSv	
	n (n_{adj})	OR (95% CI)	p	OR (95% CI)	p
Total	937	1.01 (0.97, 1.05)		0.97 (0.89, 1.06)	
Leukemia subtypes			0.20		0.28
ALL	786	1.02 (0.98, 1.07)		0.99 (0.90, 1.09)	
AML	101	0.95 (0.84, 1.08)		0.92 (0.75, 1.15)	
Other	54	1.00 (0.86, 1.17)		0.94 (0.73, 1.19)	
Age groups, years			0.005**		0.007**
2–<7	605	1.05 (1.00, 1.10)*		1.27 (1.01, 1.60)*	
7–<15	332	0.93 (0.86, 1.00)		0.93 (0.85, 1.02)	
ALL			0.10		0.22
TEL-AML1	82	0.97 (0.83, 1.12)		0.90 (0.53, 1.52)	
HeH	190 (145)	1.11 (1.02, 1.21)*		1.30 (0.94, 1.80)[1]	
Other abnormalities	233	1.01 (0.93, 1.09)		1.04 (0.89, 1.22)	
Normal	202	1.00 (0.92, 1.08)		0.96 (0.81, 1.14)	
ALL, 2–<7 years			0.30		0.50
Total	533	1.06 (1.00, 1.11)*		1.25 (0.98, 1.60)	
TEL-AML1	68	0.99 (0.85, 1.15)		1.07 (0.54, 2.13)	
HeH	147 (127)	1.16 (1.05, 1.28)**		2.00 (1.10, 3.65)[1]*	
Other abnormalities	143 (134)	1.19 (0.62, 2.11)		1.14 (0.82, 1.59)[1]	
Normal	121 (109)	1.05 (0.95, 1.17)		1.12 (0.66, 1.90)[1]	

Crude ORs were used if adjusting for Down syndrome, smoking during pregnancy and being large for gestational age (LGA) did not alter the result more than 0.05 U. Adjusted ORs are marked with [1]Statistically significant results are marked with an asterisk ($\alpha = 0.05$) and highly significant results are marked with two asterisks ($\alpha = 0.01$). All cases and controls with LGA or smoking during pregnancy missing were first excluded ($n = 774$, 203 cases and 571 controls) to make crude and adjusted ORs comparable when deciding if adjusting is necessary. If adjusting was not needed all cases and controls were used in the crude model. Shown p values represent testing for heterogeneity among adjusted coefficients in the subgroups.

significantly between the genetic abnormalities (interaction $p = 0.10$).

In exploratory analyses by age and genetic aberration, cases aged 2–<7 years with HeH genetic subtype showed the most elevated ORs (OR 2.00, 95% CI 1.10, 3.65 for 1 mSv increase in cumulative equivalent dose to RBM).

A total of 26 (2.8%) of the cases and 36 (1.3%) controls with non-zero radiation exposure had lived abroad and the uncertainty in their exposure assessment could affect the results. However, excluding these subjects did not yield different results compared with the primary analysis (OR 1.01 per 10 nSv h^{-1}, 95% CI 0.97, 1.06 for all cases) nor did adjustment for residence abroad in the regression model. Similarly, exclusion of subjects with missing data on birth weight (184 cases and 511 controls) did not alter the results for background radiation (OR 1.02 per 10 nSv h^{-1}, 95% CI 0.97, 1.06 for all cases). In addition, we approximated exposure *in utero* by assuming a 9-month exposure for each case and control in the residence at birth (Supporting Information Table S1). The effect of the 2-year minimum latency period was also approximated and Supporting Information Table S2 displays the results ignoring any latency. Neither these approaches yielded materially different results.

The potential effect of CT-imaging was evaluated in various hypothetical bias scenarios: cases receiving 20% larger dose from CT than controls (OR 1.03 per 10 nSv h^{-1}, 95% CI 0.99, 1.07), all subjects in the highest (OR 1.01 per 10 nSv h^{-1}, 95% CI 0.97, 1.05) or the lowest (OR 1.01 per 10 nSv h^{-1}, 95% CI 0.97, 1.05) quartile of background radiation receiving 20% larger doses, cases in the highest (OR 1.02 per 10 nSv h^{-1}, 95% CI 0.98, 1.06) or the lowest (OR 1.02 per 10 nSv h^{-1}, 95% CI 0.98, 1.06) quartile of background radiation receiving 20% larger dose.

Discussion

We estimated the contribution of background gamma radiation on risk for childhood leukemia based on a nationwide register-based case-control study in Finland. Overall, a small, non-significant excess risk was found for dose rate, while the point estimate for cumulative dose was below unity. However, the confidence intervals include the risk estimates obtained by extrapolation from higher dose levels, and are

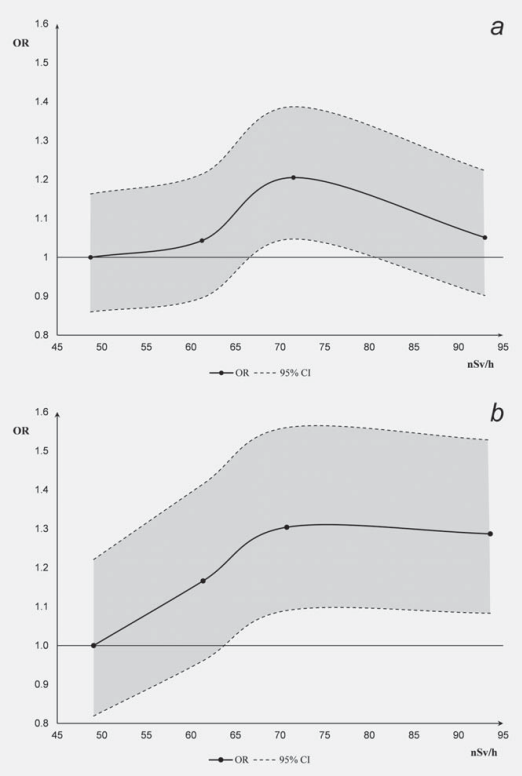


Figure 4. Dose-response relationship between increasing average equivalent dose rate and leukemia diagnoses among all cases (*a*) and among 2–<7 year age group (*b*). Cut-off values for groups are defined as quartiles of average equivalent dose rate to red bone marrow to controls. The point estimates represent medians in each group accordingly. The group formed by the lowest quartile is compared to other groups.

consistent with earlier studies of background radiation. Furthermore, a significant effect was observed for the age group 2–<7 years at diagnosis. In addition, we observed suggestive differences in OR by genetic subtype, but heterogeneity between these groups was not significant. Likewise, no significant heterogeneity was observed between different cell types (ALL, AML, others).

Our risk estimates were more stable for dose rate than cumulative dose, partly reflecting the exposure distribution (most subjects with dose rates 40–100 nSv h^{-1} and cumulative doses 0–10 mSv). Furthermore, the dose rate does not accumulate with age, though some regression towards the mean with age is expected due to increasing moves and number of residences. Therefore, in the analyses of dose rate, cases at all ages contribute with equal weight, while for cumulative doses, cases diagnosed at older ages show the highest exposures and hence have more influence on the effect at highest doses, which drive the linear risk estimates.

The anticipated effect size is small, as in the Life Span Study of atomic bomb survivors; the excess relative risk estimate for leukemia for children exposed at ages 0–9 years is of the order of 10 per Sv, corresponding to RR 1.01 per mSv.[18] Our results are largely in this range, though the confidence intervals are generally wide enough to cover ORs ranging from no effect to risks that are an order of magnitude higher. Our results are also compatible with other studies of background radiation and childhood leukemia.[6,7] The risk estimate was 1.12 per mSv (95% CI 1.03, 1.22) in the British study and 1.04 (1.00, 1.08) in the Swiss study. The highest ORs we found for the age group 2–<7 years, as well as the high hyperdiploids are well above the anticipated level, and may partly reflect chance variation.

We evaluated the risk related to background gamma radiation in relation to specific ALL subtypes, with the rationale that they represent different disease entities with potentially differing pathogenetic processes and contrasting roles for varying etiologic agents. There is little evidence to demonstrate differences in etiology between leukemia subtypes.[2,19,20]

Our study has several strengths. Our material includes a comprehensive and representative roster of all childhood leukemia cases diagnosed in Finland and the controls were identified from the population registry with complete enumeration of all residents of Finland, thus avoiding any selection bias. Unlike previous studies on the subject, we were able to construct complete residential histories for all subjects including both place of residence and type of dwelling, with register-based data eliminating recall bias. For exposure assessment, we compiled dose rates from both natural sources and the transient increase due to the Chernobyl fallout. For converting outdoor to indoor dose rates, we used experimental conversion factors by type of dwelling, taking into account the effect of building materials, while also factoring in occupancy of Finnish children and varying body composition by age during childhood. Our results also showed the effects of known risk factors for childhood leukemia, reaffirming the validity of the data and minimizing confounding. Furthermore, this is the first study in exploring the role of background radiation in relation to various genetic subtypes of leukemia.

Some shortcomings may limit our ability to precisely capture the effect of background radiation on leukemia risk. Long-term individual-level dose monitoring would provide the ideal exposure estimates, but it is not feasible. Reconstructing exposure from surveys providing estimates by geographic area induces inevitably some misclassification. Some variation can be expected to occur even between similar dwelling types within an 8 km × 8 km square used as a geographical unit in our analysis, and with children with contrasting occupancy. Such misclassification can be expected to be nondifferential, *i.e.*, affect the cases and controls similarly, and hence attenuate the observed effect from the true one. This very likely also applies to radiation exposure from other sources, notably medical diagnostic radiation. The mean

annual effective dose from medical radiation in Finland is 0.45 mSv (2012), but for children it is markedly lower due to strict imaging guidelines especially with CT scans, which have the largest contribution to the annual dose from medical exposure.[8,21] Also, CT usage on Finland is less frequent than many other industrialised countries and only 1.7% of CT scans in Finland were performed on children.[22] Hence, medical radiation exposure is relatively small compared to annual doses observed in our study (0.59 mSv to RBM). In addition, CT-imaging would cause confounding only if it is associated with natural background gamma radiation exposure.

We evaluated the potential impact of CT examinations through quantitative bias analyses. We obtained estimates of effective dose from the literature[23] and computed average annual doses using data on the use of diagnostic imaging obtained from Finnish radiation and nuclear safety authority.[22] The average effective dose per CT examination in Finland was 2.75 mSv, corresponding to an annual average effective dose of 0.02 mSv per child, which is minimal compared to background radiation. We also evaluated various hypothetical bias scenarios: cases receiving 20% larger dose from CT than controls, all subjects in the highest or the lowest quartile of background radiation receiving 20% larger doses, cases in the highest or the lowest quartile of background radiation receiving 20% larger dose. Overall, these analyses revealed no material confounding indicating the robustness of our estimates in relation to CT examinations.

The statistical power was relatively low, given the small expected effect size. The minimum sample size for studies of natural background radiation has been calculated as 7,800 cases with five times as many controls to reach a statistical power of 80%.[24] These power calculations do not, however, take into account the precision and validity of the exposure assessment, which is also an important determinant of the amount of information that a study can contribute. For our study, improvement in these aspects is the most important addition to the evidence, compared with earlier publications.

Conclusions

Overall, we found no significant increase in OR of childhood leukemia with dose or dose rate from background radiation. However, in subgroup analyses, we observed a significantly increased OR in relation to both dose rate and cumulative dose in the agegroup 2–<7 years. In addition, leukemia with high hyperdiploidy showed a significant association with background gamma radiation, unlike other genetic subtypes, but this requires further confirmation.

Acknowledgements

We thank clinicians Kim Vettenranta (Helsinki), Päivi Lähteenmäki (Turku), Merja Möttönen (Oulu) and Jouni Pesola (Kuopio) for collecting the genetic subgroup data. We are grateful for the insightful comments by Kajsa Paulsson and Kaj Vesterbacka.

References

1. Inaba H, Greaves M, Mullighan CG. Acute lymphoblastic leukaemia. *Lancet* 2013;381:1943–55.
2. Hsu WL, Preston DL, Soda M, et al. The incidence of leukemia, lymphoma and multiple myeloma among atomic bomb survivors: 1950–2001. *Radiat Res* 2013;179:361–82.
3. Hasle H. Pattern of malignant disorders in individuals with Down's syndrome. *Lancet Oncol* 2001;2:429–36.
4. Roman E, Simpson J, Ansell P, et al. Perinatal and reproductive factors: A report on haematological malignancies from the UKCCS. *Eur J Cancer* 2005;41:749–59.
5. Roberts I, Izraeli S. Haematopoietic development and leukaemia in Down syndrome. *Br J Haematol* 2014;167:587–99.
6. Kendall GM, Little MP, Wakeford R, et al. A record-based case-control study of natural background radiation and the incidence of childhood leukaemia and other cancers in Great Britain during 1980–2006. *Leukemia* 2013;27:3–9.
7. Spycher BD, Lupatsch JE, Zwahlen M, et al. Background ionizing radiation and the risk of childhood cancer: a census-based nationwide cohort study. *Environ Health Perspect* 2015;123:622–8.
8. UNSCEAR. United Nations Scientific Committee on the Effects of Atomic Radiation 2008 Report to the General Assembly, Vol. 1. Effects of Ionising Radiation with Scientific Annexes. 2010; United Nations: New York.
9. Sankilampi U, Hannila ML, Saari A, et al. New population-based references for birth weight, length, and head circumference in singletons and twins from 23 to 43 gestation weeks. *Ann Med* 2013;45:446–54.
10. Arvela H, Hyvönen H, Lemmelä H, et al. Indoor and outdoor gamma radiation in Finland. *Radiation Protect Dosimetr* 1995;59:25–32.
11. Arvela H, Markkanen M, Lemmelä H. Mobile survey of environmental gamma radiation and fall-out levels in Finland after the chernobyl accident. *Radiat Protect Dosimetr* 1990;32: 177–84.
12. Mäkeläinen I, Moisio S, Reisbacka H, et al. Indoor occupancy and radon exposure in Finland. *Radioactiv Environ* 2005;7:687–93.
13. Saito K, Petoussi N, Jacob P, et al. Calculation of organ doses from environmental gamma rays using human phantoms and Monte Carlo methods: part 1 monoenergetic sources in the ground and natural radionuclides in the ground. *GSF Bericht* 1990;2:90.
14. Petoussi N, Jacob P, Zankl M, et al. Organ doses for foetuses, babies, children and adults from environmental gamma rays. *Radiat Protect Dosimetr* 1991;37:31–41.
15. Kendall GM, Fell TP, Harrison JD. Dose to red bone marrow of infants, children and adults from radiation of natural origin. *J Radiol Prot* 2009;29: 123–38.
16. Easton DF, Peto J, Babiker AG. Floating absolute risk: an alternative to relative risk in survival and case-control analysis avoiding an arbitrary reference group. *Stat Med* 1991;10:1025–35.
17. Lachin JM. Sample size evaluation for a multiply matched case-control study using the score test from a conditional logistic (discrete Cox PH) regression model. *Stat Med* 2008;27:2509–23.
18. Preston DL, Kusumi S, Tomonaga M, et al. Cancer incidence in atomic bomb survivors. Part III. Leukemia, lymphoma and multiple myeloma, 1950–1987. *Radiat Res* 1994;137: S68–97.
19. Sandler ES, Friedman DJ, Mustafa MM, et al. Treatment of children with epipodophyllotoxin-induced secondary acute myeloid leukemia. *Cancer* 1997;79:1049–54.
20. Weiss B, Vora A, Huberty J, et al. Secondary myelodysplastic syndrome and leukemia following 131I-metaiodobenzylguanidine therapy for relapsed neuroblastoma. *J Pediatr Hematol Oncol* 2003;25:543–7.
21. Muikku M, Bly R, Kurttio P, et al. The average effective dose in Finland, in Finnish. Radiation and Nuclear Safety Authority 2014; STUK-A259.
22. Helasvuo T. The number of radiological imaging investigations and operations in the year 2011, in Finnish. Radiation and Nuclear Safety Authority 2013;STUK-B161.
23. Chapple CL, Willis S, Frame J. Effective dose in paediatric computed tomography. *Phys Med Biol* 2002;47:107–15.
24. Little MP, Wakeford R, Lubin JH, et al. The statistical power of epidemiological studies analyzing the relationship between exposure to ionizing radiation and cancer, with special reference to childhood leukemia and natural background radiation. *Radiat Res* 2010;174:387–402.

PUBLICATION II

Effects of incomplete residential histories on studies of environmental exposure with application to childhood leukemia and background radiation

Atte Nikkilä, Gerald Kendall, Jani Raitanen, Ben Spycher, Olli Lohi, Anssi Auvinen

Environmental Research 166, 466-472
https://doi.org/10.1016/j.envres.2018.06.035

Environmental Research 166 (2018) 466–472

Contents lists available at ScienceDirect

Environmental Research

journal homepage: www.elsevier.com/locate/envres

Effects of incomplete residential histories on studies of environmental exposure with application to childhood leukaemia and background radiation

Atte Nikkilä[a,*,1], Gerald Kendall[b,1], Jani Raitanen[c,d], Ben Spycher[e], Olli Lohi[f], Anssi Auvinen[c,g]

[a] Faculty of Medicine and Biosciences, University of Tampere, Tampere, Finland
[b] Cancer Epidemiology Unit, NDPH, University of Oxford, Old Road Campus, Oxford OX3 7LF, UK
[c] Faculty of Social Sciences, University of Tampere, Tampere, Finland
[d] UKK Institute, Tampere, Finland
[e] Institute of Social and Preventive Medicine (ISPM), University of Bern, Bern, Switzerland
[f] Tampere Center for Child Health Research, University of Tampere and Tampere University Hospital, Tampere, Finland
[g] Radiation and Nuclear Safety Authority, Helsinki, Finland

ARTICLE INFO

Keywords:
Residential history
Exposure assessment
Childhood leukaemia

ABSTRACT

When evaluating environmental exposures, residential exposures are often most relevant. In most countries, it is impossible to establish full residential histories. In recent publications, childhood leukaemia and background radiation have been studied with and without full residential histories. This paper investigates the consequences of lacking such full data.

Data from a nationwide Finnish Case-Control study of Childhood Leukaemia and gamma rays were analysed. This included 1093 children diagnosed with leukaemia in Finland in 1990–2011. Each case was matched by gender and year of birth to three controls. Full residential histories were available. The dose estimates were based on outdoor background radiation measurements. The indoor dose rates were obtained with a dwelling type specific conversion coefficient and the individual time-weighted mean red bone marrow dose rates were calculated using age-specific indoor occupancy and the age and gender of the child. Radiation from Chernobyl fallout was included and a 2-year latency period assumed.

The median separation between successive dwellings was 3.4 km and median difference in red bone marrow dose 2.9 nSv/h. The Pearson correlation between the indoor red bone marrow dose rates of successive dwellings was 0.62 (95% CI 0.60, 0.64). The odds ratio for a 10 nSv/h increase in dose rate with full residential histories was 1.01 (95% CI 0.97, 1.05). Similar odds ratios were calculated with dose rates based on only the first dwelling (1.02, 95% CI 0.99, 1.05) and only the last dwelling (1.00, 95% CI 0.98, 1.03) and for subjects who had lived only in a single dwelling (1.05, 95% CI 0.98, 1.10).

Knowledge of full residential histories would always be the option of choice. However, due to the strong correlation between exposure estimates in successive dwellings and the uncertainty about the most relevant exposure period, estimation of overall exposure level from a single address is also informative. Error in dose estimation is likely to cause some degree of classical measurement error resulting in bias towards the null.

1. Introduction

Environmental exposure is often determined by location. Because children typically spend most of their time at home (UNSCEAR, 2000), residential exposures, also known as domestic exposures, are often most relevant. Examples include residential exposure to background radiation for residential radon and terrestrial gamma rays, electromagnetic fields and air pollution (Kroll et al., 2010; Raaschou-Nielsen et al., 2001; UK Childhood Cancer Study Investigators, 1999).

Children are more susceptible to the carcinogenic effects of ionising radiation than adults, and many studies of natural radiation and cancer have focused on children (Demoury et al., 2017; Kendall et al., 2013; Nikkilä et al., 2016; Raaschou-Nielsen et al., 2008; Spix et al., 2017; Spycher et al., 2015). Such studies must include thousands of cases in order to achieve sufficient statistical power to detect the small effects expected from ambient exposures, as extrapolated from high dose levels (Little et al., 2010). However collecting information on such large samples through interview or direct measurement is expensive and prone to selection bias. In most

* Corresponding author.
E-mail address: atte.nikkila@uta.fi (A. Nikkilä).
[1] Equal contributions.

https://doi.org/10.1016/j.envres.2018.06.035
Received 14 March 2018; Received in revised form 13 June 2018; Accepted 16 June 2018

countries, it is impossible to establish full residential histories, i.e. a list of dwellings occupied with dates of moving in and out, without individual contact with study subjects. Only few countries, including the Nordic countries, have nationwide registries that provide such data. For this reason, many such studies have limited exposure assessment to a single dwelling, for example that occupied at diagnosis (UK Childhood Cancer Study Investigators, 2002a, 2002b; Demoury, 2017) or that at birth (Kendall et al., 2013).

Because direct measures of natural radiation are impracticable, exposure is usually estimated from models. The indoor dose rates of terrestrial gamma radiation depend on outdoor dose rates, the shielding effect of the material of the house, and the radiation emitted by the building materials. Reasonable predictions of indoor gamma-ray dose rates can be made based on location and factors such as geology and socioeconomic status of the area (Chernyavskiy et al., 2016; Kendall et al., 2016; Warnery et al., 2015). Nevertheless, considerable inter-house variation remains. In the British studies, the residual Mean Square Error (MSE) was 378 $(nGy/h)^2$ on a mean estimate of 96 nGy/h; in France MSE = 407 $(nSv/h)^2$ with a mean of 76 nSv/h. Such significant uncertainties can be avoided only by direct measurement in the house in question or, potentially, by modelling using more detailed information, in particular on the radioactive content of all significant building materials used in each dwelling (European Commission, 1997). Similar considerations apply, perhaps with even more force, to indoor radon concentrations – neighbouring houses can differ greatly in a random, time varying and unpredictable way, but models can make reasonable predictions for areal averages (Miles and Appleton, 2005).

In studies of the effects of protracted exposures to ionising radiation (or other agents) it is necessary to consider how susceptibility varies over the exposure period. This is particularly important in the context of background radiation, as doses are accumulated throughout life and children are known to be more susceptible (UNSCEAR, 2008a). The extent to which susceptibility varies throughout childhood is unknown though there is evidence that susceptibility is greatest at younger ages (National Research Council NRC, 2006; UNSCEAR, 2008a, 2013). A reasonable approach is to use total cumulative dose or time-integrated dose rates from birth or conception to diagnosis as exposure measure. However, it is possible that, for example, exposures around the time of birth or during pregnancy are disproportionately important (ICRP: International Commission on Radiological Protection, 2003; National Research Council NRC, 2006; UNSCEAR, 2013).

A recent publication from the Finnish Register-based Case-Control study of Childhood Leukaemia (FRECCLE) (Nikkilä et al., 2016), on the effect of natural gamma rays on risk of childhood leukaemia made use of a nationwide population registry in order to obtain complete residential histories for the study subjects. The present paper analyses these data in more detail to obtain insights for the interpretation of similar epidemiological studies, which lack the richness of the data available in Finland.

It is the aim of this paper

i) to investigate patterns of residential mobility, or migration, in families with young children,
ii) to investigate the extent to which doses in successive dwellings vary and
iii) to examine the implications of residential mobility for epidemiological studies of natural gamma rays and childhood cancers in general and leukaemia in particular.

2. Materials and methods

FRECCLE includes 1093 children diagnosed with leukaemia in Finland in 1990–2011, identified from the Finnish Cancer Registry. Each case was matched by gender and year of birth to three controls at the Population Register Centre. For the matched cases and controls, the Population Register Centre provided complete residential histories from birth to the reference date (date of diagnosis for cases; for controls, date when an exposure period of similar length is reached). These residential histories included moving dates and municipalities of residence, as well as coordinates, building type and identification codes for each residence.

The ambient dose estimates are based on an 8 × 8 km gridded map of outdoor natural background radiation from the Finnish Radiation and Nuclear Safety Authority. This map is based on the measurements from a mobile survey carried out between 1978 and 1980 in Finland (Arvela et al., 1995). The indoor dose rates were obtained with a conversion coefficient specific to the dwelling type and the individual time-weighted red bone marrow dose rate averages were calculated using age-specific indoor occupancy coefficients and the age and gender of the child (Arvela et al., 1995; Kendall et al., 2009; Mäkeläinen et al., 2005). Radiation from Chernobyl fallout was also modelled, though its contribution to the total dose estimates was small (~3%) (Nikkilä et al., 2016). The median dose rate to red bone marrow (Chernobyl and natural background radiation) was 67.2 nSv/h for cases and 66.4 nSv/h for controls.

Based on previous studies, Nikkilä et al. (2016) assumed a 2-year minimum latency period in their main analysis resulting in zero exposure for subjects younger than 2 years at the reference date (in utero exposure was ignored) (UNSCEAR, 2008b). The odds ratio (OR) of childhood leukaemia was calculated for every 10 nSv/h increase in time-weighted average indoor gamma-ray dose-rate over the period from birth to the reference date. For the present work other measures of exposure were investigated as described in the Results section.

Statistical analyses were done using R (3.4.0) and conditional logistic regression was used for matched case-control data. The ethical committee of Pirkanmaa Hospital district reviewed the study protocol (tracking number R14074). According to Finnish regulations, no informed consent was required for a register-based study.

3. Results

Table 1 shows the number of dwellings occupied by cases and by controls in the Finnish study (Nikkilä et al., 2016). These numbers take into account the two-year latency period, so that subjects aged less than 2 years at their reference date (156 cases and 468 controls) were excluded, leaving 937 cases and 2811 controls for analysis. About 48% of both cases and controls had lived at only one address between birth and the reference date. Five percent of both cases and controls had lived in five or more dwellings during the exposure period. The mean number of addresses occupied between birth and the reference date was approximately 1.9 for both cases and controls. In total, there were 63 (0.8%) residencies abroad. The percentage of cases who moved during the two-year latency period preceding the reference date was 26.0% and 26.2% for controls.

Table 2 shows the separation (km) and mean difference in dose rate for different pairs of addresses for cases and controls separately. The first of these is equivalent to the separation of two randomly chosen dwellings in our Finnish dataset; the median is similar to the mean for cases and controls (cases: 233 km vs. 267 km, controls: 230 km vs. 264 km). Successive homes of the same family are, on average, much closer together, with a median separation of only 3.4 km (cases: 3.6 km, controls: 3.3 km). The mean distance between successive dwellings of the same family is an order of magnitude larger than the median, because of the influence of relatively uncommon long-distance moves (i.e. the distribution is highly skewed). The separation of the first and last dwellings occupied by a family is a little larger than the separation of successive homes, but broadly similar. The dose rates varied as might be expected, with successive dwellings of the same family having the lowest changes in dose rate.

The Pearson correlation coefficients with their 95% confidence intervals for four selected pairs of dose-rate variables (“scenarios”) are presented in the Table 3. For all scenarios, the indoor and outdoor dose rates for all study subjects were analysed separately. The correlation coefficient between indoor gamma dose rates for two successive dwellings occupied by study subjects was 0.62 (95% CI 0.60, 0.64). The correlation between the dose rates at the first and the last dwelling occupied by the study subjects was 0.67 (95% CI 0.65, 0.70). The correlation between the indoor dose rate averaged over all the subject's dwellings and the dose rate inside the first dwelling was 0.78 (95% CI 0.76, 0.80). The respective values of these correlation coefficients for outdoor dose rates were higher. The Pearson

Table 1
Number of dwellings occupied by study subjects.

	1	2	3	4	5	≥ 6	Total	max
Controls	1459 (52%)	721 (26%)	342 (12%)	150 (5%)	72 (3%)	67 (2%)	2811	11
Cases	488 (52%)	257 (27%)	103 (11%)	44 (5%)	25 (3%)	20 (2%)	937	16

Note that account has been taken of a two year latency period in these figures.
Note that the data of Table 1 are reproduced by permission from the International Journal of Cancer.

Table 2
Mean separation (km) and mean absolute difference in dose rate (nSv/h) between two random addresses and between addresses for the same family for cases and controls.

	Median	Mean	Q1	Q3	Max
Any two addresses[a]					
Separation (km)					
Cases	233	267	135	371	1144
Controls	230	264	133	370	1179
Absolute difference in dose rate (nSv/h)					
Cases	14.5	17.2	6.2	25.5	102
Controls	14.4	17.4	6.1	26.0	102
Successive addresses					
Separation (km)					
Cases	3.6	36.0	1.0	12.1	792
Controls	3.3	29.9	0.9	13.6	760
Absolute difference in dose rate (nSv/h)					
Cases	2.4	8.7	0	16.5	96.1
Controls	3.3	9.2	0	17.7	78.7
First and last address					
Separation (km)					
Cases	4.3	37.2	1.16	15.0	713
Controls	4.2	33.3	1.19	16.3	760
Absolute difference in dose rate (nSv/h)					
Cases	4.7	8.9	0	17.0	39.7
Controls	3.8	9.0	0	17.4	74.5
Difference in dose rate (nSv/h)[b]					
Cases	− 0.1	− 4.6	− 14.5	0	39.5
Controls	− 0.1	− 5.1	− 15.3	0	47.1

[a] All pairs of addresses in the dataset, not necessarily from the same subject.
[b] The difference is calculated by substracting the dose rate in the last dwelling from the dose rate in the first dwelling.

Table 3
Pearson correlation coefficients with 95% confidence intervals between various In- and Outdoor gamma-ray dose rate quantities.

	Indoors		***Outdoors***	
	r	***95% CI***	***r***	***95% CI***
Dose rates in successive dwellings				
Cases	0.59	0.55–0.64	0.85	0.83–0.87
Controls	0.64	0.62–0.67	0.89	0.88–0.89
Dose rates in first and last dwelling[a]				
Cases	0.67	0.61–0.71	0.91	0–89–0.92
Controls	0.67	0.64–0.70	0.88	0.87–0.89
Dose rate in first vs mean for all dwellings				
Cases	0.77	0.73–0.81	0.99	0.98–0.99
Controls	0.78	0.76–0.80	0.98	0.98–0.98
Change in dose rate vs separation of successive dwellings				
Cases	0.31	0.24–0.37	0.68	0.63–0.71
Controls	0.22	0.18–0.26	0.60	0.58–0.63

[a] If only subjects with three or more dwellings are included, the correlation for indoor values lower for cases (0.54, (0.43, 0.63)) than for controls (0.63 (0.58, 0.68)).

correlation between the separation and difference in the indoor effective dose rate for pairs of successive dwellings was quite low (0.32, 95% CI 0.29, 0.35). In all these analyses, results for cases were similar to those for controls.

Nikkilä et al. (2016) reported that the odds ratio of childhood leukaemia for every 10 nSv/h increase in dose rate based on the time-weighted average indoor gamma-ray dose rate over all residences in the time window was 1.01 (95% CI 0.97, 1.05). The corresponding OR, based on the dose rate in only the first and the last dwelling, are 1.02 (95% CI 0.99, 1.05) and 1.00 (95% CI 0.98, 1.03), respectively, for every 10 nSv/h increase. When the subgroup of subjects with only one residence in their history (488 cases and 1458 controls) was analysed, we observed an OR of 1.05 (95% CI 0.98, 1.11) for every 10 nSv/h increase. When we modelled the indoor dose rates naively as the outdoor dose rates, neglecting the effects of the dwelling type, we observed an OR of 1.04 (95% CI 1.00, 1.08) for every 10 nSv/h increase using the full residential histories.

When the cumulative RBM dose is considered, with full residential histories an OR of 0.97 (95% CI 0.89, 1.06) per 1 mSv was previously reported (Nikkilä et al., 2016). In analyses considering only the first and only the last dwelling the ORs per 1 mSv were respectively: 1.04 (95% CI 0.99, 1.10) and 1.00 (95% CI 0.95, 1.06). For the subset of subjects with only one dwelling in their history we observed OR of 1.20 (95% CI 0.94, 1.52) and if the analysis is completed based on outdoor dose rates with full residential histories an OR of 0.98 (95% CI 0.90, 1.08) was observed. The results based on average dose rate and cumulative RBM dose with a two-year latency period are presented in Fig. 1.

When the more mobile subjects, with at least two dwellings in their residential history (425 cases and 1304 controls), were analysed separately we observed an OR of 0.97 (95% CI 0.90, 1.05) for every 10 nSv/h increase. To explore the effect of our chosen latency period, we calculated odds ratios for cumulative RBM for no latency period and for a latency period of five years (Supplementary table 1).

4. Discussion

Residential mobility, that is the movement of study subjects from one residential address to another, is a challenge for epidemiological studies, where exposure to the agent of interest depends on place of residence. The data of the Finnish study (Nikkilä et al., 2016) demonstrate that, in their study subjects, about half the children diagnosed with leukaemia had not moved from the birth address by the time of diagnosis. Those who had moved generally had not moved far; the median separation of successive dwellings being 3.4 km.

A recent paper (Kendall et al., 2015) examined residential mobility in a case-control study, in which addresses were available for both the residence at birth and the residence at diagnosis of cases. Kendall et al. reported that 44% of all cases of childhood cancer had not moved by the time of diagnosis. For leukaemias, the figure was 45%. Those who had moved to a new house did not, on average, move far: median distance 3.1 km for both childhood leukaemia and all childhood cancers. In consequence, the estimated indoor gamma ray dose rates in the house at birth were strongly correlated with those in the house at diagnosis. These results are consistent with those from the Finnish data.

In a recent nationwide study on childhood cancer and background radiation in Switzerland (Spycher et al., 2015), residential locations were available for all children at the time of national censuses. In analyses using cumulative doses since birth, the hazard ratio for childhood leukaemia was higher in a sub-cohort with stable place of residence for at least 5 years before census (excess relative risk per mSv: 4.6%, 95% CI: −0.1–9.6%) compared to the full cohort (3.6%, −0.3–7.7%), in which mobility was

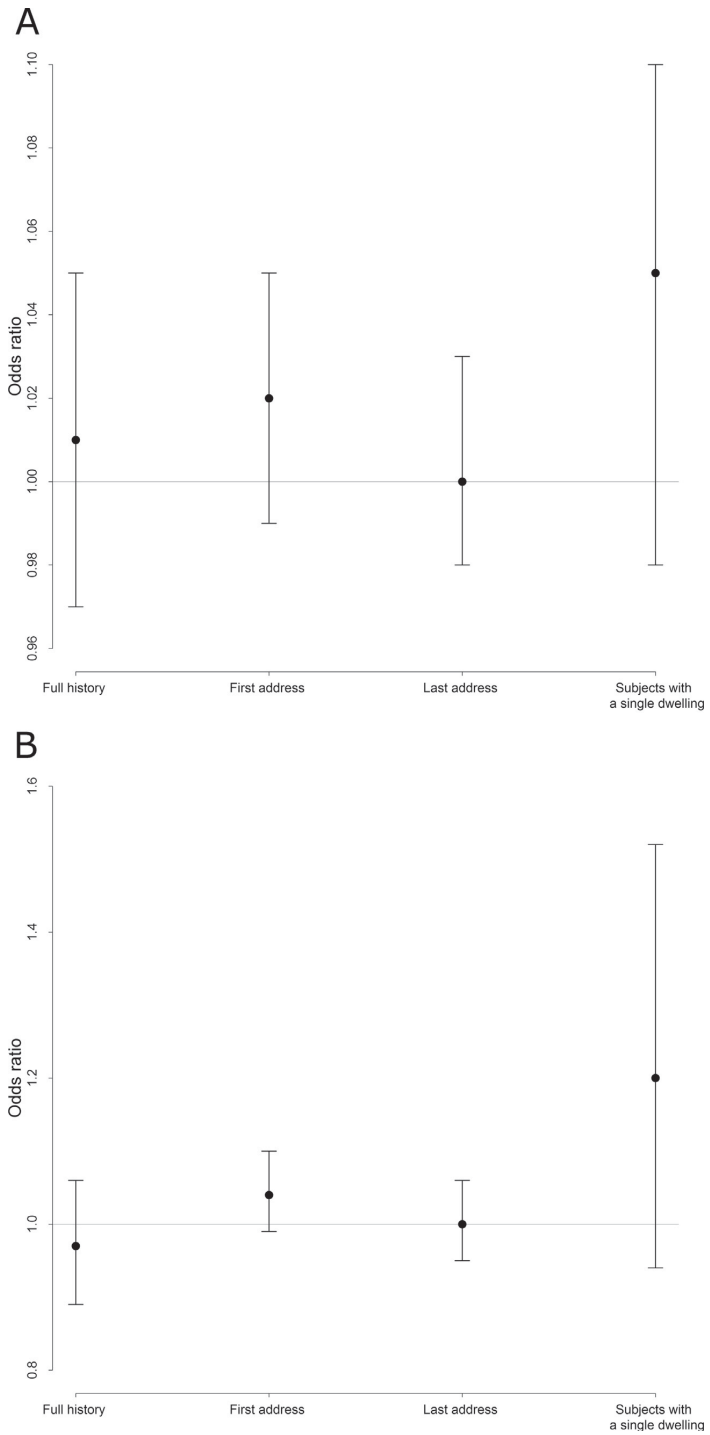


Fig. 1. Odds ratios for 10 nSv/h increase in average red bone marrow dose rate (A) and for 1 mSv increase in cumulative red bone marrow dose (B) for childhood leukaemia and terrestrial background radiation (width: 1 column). Different scenarios were chosen to approximate exposure assessment based on residential histories. Full history: Full residential histories. First address: Only first dwelling was included. Last address: Only last dwelling was included. Subjects with a single dwelling: Only subjects with one residence in their history were included.

more common. For cases, full residential histories were available from the national childhood cancer registry. These data show that 34% of leukaemia cases moved house at least once between birth and diagnosis (Kreis et al., 2016), somewhat fewer than in the present study.

Demoury et al. (2017) note that in the ESCALE interview-based study of childhood cancers and leukaemia in France, 66% of the children had been living in the same municipality (Commune) since birth. The correlations between exposure estimates at birth and at diagnosis (cases) or inclusion (controls) were 0.86 for radon exposure and 0.89 for gamma radiation exposure. A Danish study by Raaschou-Nielsen et al. (2008) reported that in their data 58% of the families had lived in a single-family dwelling throughout their childhood.

In their published paper, Nikkilä et al. (2016) chose as their principal exposure measure the indoor gamma-ray dose rate averaged over all dwellings. Here, the effect of not having complete residential histories is explored by using several other exposure indicators (based on only first dwelling, only last dwelling, subgroup of subjects with only one dwelling and dose-rate estimated neglecting the dwelling type). The confidence intervals of the ORs overlap markedly with results for the average dose rate, though the central estimates ranged from 1.00 to 1.05, while the published result was 1.01. A similar picture is seen for results based on cumulative dose for which the point estimates were more dispersed: from 0.97 to 1.20. The sensitivity analyses with differing length latency periods reflected the previously reported results – in our material, the age of the subject appears to be an effect modifier (higher ORs were observed for younger children).

Patterns of migration may vary from country to country and from time to time. Further, exposure difference between successive homes may depend on the size of the country and spatial variation of dose rates. Childhood leukaemia peaks at a younger age than most other paediatric cancers and therefore residential histories are likely to be somewhat shorter and involve fewer residencies (although Kendall et al. (2015) reported that such differences were not great). Hence, there will always be uncertainties in extrapolating from one set of circumstances to others.

Interestingly, it has been hypothesized that a particular migration pattern characterised by rapid population influx into isolated rural areas might play a role in the aetiology of childhood leukaemia (Kinlen et al., 2012). Hence, in such specific instances, moving might not only induce exposure measurement error, but also confounding. Thus, having complete residential histories might enhance control of confounding besides producing more accurate estimates of exposure.

Considerable uncertainty remains about the most important time window for leukaemogenesis in childhood. If study subjects move house between birth and the start of the latency period then adequate residential histories are needed to establish the precise mean dose rate or integrated dose over this exposure period. However, if susceptibility varies with age, such time-weighted averages based on subjects' whole residential histories are not necessarily the most relevant. Different weights should perhaps be given to the birth address for example (ICRP: International Commission on Radiological Protection, 2003; National Research Council NRC, 2006; UNSCEAR, 2013). If latency plays an important role, modelling exposure by the dwelling at diagnosis is not optimal. That is because a proportion (in our dataset a fourth, assuming two-year latency) of study subjects will have moved during their latency period. For these subjects, there would be no relevant information available about their exposure during the (presumed) etiologically relevant time window. Of course, with full residential histories it is theoretically possible to explore sensitivity during different time windows using the data themselves. However, this would require a larger study with greatly increased statistical power.

The ORs observed in the present study were somewhat higher for cases who had lived in only a single dwelling in history compared to those with two or more. This reflects the previously reported difference in leukaemia risk by age group as the mean number of dwellings increases with age.

With dose metrics averaged over the exposure periods of subjects, such as the average dose rate, regression towards mean with an increasing number of dwellings in history will occur, and it is likely to decrease variation between subjects and hence diminish exposure contrast and dilute the extremes. However, the extent of regression toward mean depends correlation between successive dwellings which was found to be high.

How much will exposures in successive dwellings vary? Clearly this depends on the agent in question. However, for natural gamma rays there is good evidence that dose rates in adjacent locations tend to be similar. In consequence, indoor gamma-ray doses in successive dwellings are highly correlated. This is supported by the present study and by Kendall et al. (2015) who reported, for Great Britain, a correlation in estimated indoor gamma rate between birth and diagnosis locations of 0.48 for cases who had moved County District and of 0.90 for all cases (Kendall et al., 2015). However, the reasons for this correlation may be complex; the data of Table 3 suggest that it is not just a proximity effect as the correlation between separation and difference in effective indoor dose rate was not as high as was expected.

So far as other exposures are concerned, our results should be generalised only with caution. It will remain true that successive family dwellings will tend to be close together. It is plausible that exposures to many agents of interest will therefore tend to be similar, but this judgement must be made on a case-by-case basis.

5. Summary

Knowledge of full residential histories would always be the option of choice. If dose estimates are constructed for the full exposure periods, having full residential histories results in less exposure misclassification. However, there is considerable evidence that families with young children who move house generally do not move far. In consequence exposures to indoor gamma rays (and possibly to other agents) in successive homes are correlated. Given this strong correlation between exposure estimates in successive dwellings and the uncertainty about the most relevant exposure period, estimation of overall exposure level from a single address, in particular that at birth, will also be informative. It remains true that comparisons of results across studies using different time points of exposure (e.g. birth and diagnosis) could indicate periods of increased susceptibility provided that the studies have adequate power. As noted in the introduction there is often substantial random variation between practical dose estimates from modelling and the (unknown) true residential doses. Thus, there is likely to be a degree of classical measurement error in the dose estimates, which will bias risk estimates towards the null.

Acknowledgements

We thank Hannu Arvela for his contributions in collecting the Finnish background radiation data and helping with the dose estimation. We are grateful to John Harrison and Richard Wakeford for helpful discussions.

Conflict of interest

The authors have no conflicting interests to disclose.

Funding

The work of AN was supported by Väre Foundation for Pediatric Research, Finnish Foundation for Pediatric Research and Competitive State Research Financing of the Expert Responsibility area of Tampere University Hospital (9T030 and 9U030).

The work of GMK was supported by Children with Cancer (UK) under grant number HTRVVK0.

The review of the study protocol by the local ethical committee

Scanned documents

TAMPEREEN YLIOPISTOLLISEN SAIRAALAN ERITYISVASTUUALUEEN ALUEELLINEN EETTINEN TOIMIKUNTA

OTE PÖYTÄKIRJASTA Nro 5/2014

28.5.2014

KOKOUSTIEDOT

Aika 27.5.2014 klo 12-13.50
Paikka FM 5, Tiedekeskuksen neuv.huon

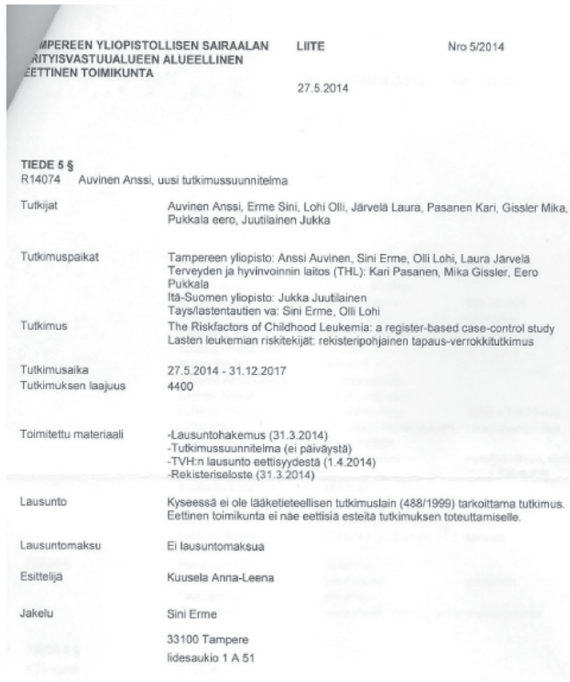
OSALLISTUJAT			
	Pasternack Amos	LKT, professori (emer.)	puheenjohtaja
	Alanen Seija	TtT, vs. ylihoitaja	
	Arvola Pertti	apulaisylilääkäri	läsnä 3-42§
	Joima Pasi	erikoislääkäri	varajäsen
	Korhonen Riku	kliininen opettaja	
	Kotisaari Marja-Liisa	OTL/VT	
	Kujala Paula	ylilääkäri	
	Kuusela Anna-Leena	apulaisylilääkäri	
	Laitinen Minna	apulaisylilääkäri	
	Lehtinen Klaus	toimialuejohtaja	läsnä 4,7-8,39-42§
	Lehto Matti	professori, lääket.yks.joht.	varapuheenjohtaja
	Mattila Elina	TtT, ylihoitaja	
	Paarlahti Teemu	sairaalapastori	maallikkojäsen, varajäsen
	Saarinen Jukka	apulaisylilääkäri	läsnä 1-4, 8-42§
	Snellman Erna	ylilääkäri	
	Tiainen Virpi	toimittaja	maallikkojäsen
	Kohonen Kirsi	TtM,palveluesimies	sihteeri
	Lahtinen Minna	TtK,tutkimussihteeri	sihteeri
POISSA	Hietaharju Aki	osastonylilääkäri	
	Nuolivirta Kirsi	erikoislääkäri	varajäsen
	Saarinen Kari	ylilääkäri	
	Sipola Antti	sairaalasielunhoidon johtaj	maallikkojäsen

TIEDE 5 §

ETL-koodi	R14074
Tutkimus	Auvinen Anssi, uusi tutkimussuunnitelma
Lausunto	Kyseessä ei ole lääketieteellisen tutkimuslain (488/1999) tarkoittama tutkimus. Eettinen toimikunta ei näe eettisiä esteitä tutkimuksen toteuttamiselle.

Pöytäkirjan otteen oikeaksi todistaa

Kirsi Korhonen
sihteeri



MPEREEN YLIOPISTOLLISEN SAIRAALAN RITYISVASTUUALUEEN ALUEELLINEN ETTINEN TOIMIKUNTA

LIITE Nro 5/2014

27.5.2014

TIEDE 5 §
R14074 Auvinen Anssi, uusi tutkimussuunnitelma

Tutkijat	Auvinen Anssi, Erme Sini, Lohi Olli, Järvelä Laura, Pasanen Kari, Gissler Mika, Pukkala eero, Juutilainen Jukka
Tutkimuspaikat	Tampereen yliopisto: Anssi Auvinen, Sini Erme, Olli Lohi, Laura Järvelä Terveyden ja hyvinvoinnin laitos (THL): Kari Pasanen, Mika Gissler, Eero Pukkala Itä-Suomen yliopisto: Jukka Juutilainen Tays/lastentautien va: Sini Erme, Olli Lohi
Tutkimus	The Riskfactors of Childhood Leukemia: a register-based case-control study Lasten leukemian riskitekijät: rekisteripohjainen tapaus-verrokkitutkimus
Tutkimusaika	27.5.2014 - 31.12.2017
Tutkimuksen laajuus	4400
Toimitettu materiaali	-Lausuntohakemus (31.3.2014) -Tutkimussuunnitelma (ei päiväystä) -TVH:n lausunto eettisyydestä (1.4.2014) -Rekisteriseloste (31.3.2014)
Lausunto	Kyseessä ei ole lääketieteellisen tutkimuslain (488/1999) tarkoittama tutkimus. Eettinen toimikunta ei näe eettisiä esteitä tutkimuksen toteuttamiselle.
Lausuntomaksu	Ei lausuntomaksua
Esittelijä	Kuusela Anna-Leena
Jakelu	Sini Erme 33100 Tampere lidesaukio 1 A 51

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.envres.2018.06.035.

References

Arvela, H., et al., 1995. Indoor and Outdoor Gamma Radiation in Finland. Radiat. Prot. Dosim. 59, 25–32.

Chernyavskiy, P., et al., 2016. Spatial prediction of naturally occurring gamma radiation in Great Britain. J. Environ. Radioact. 164, 300–311.

Demoury, C., et al., 2017. Residential exposure to natural background radiation and risk of childhood acute leukemia in France, 1990–2009. Environ. Health Perspect. 125, 714–720.

European Commission, 1997. Radiation protection 96 Enhanced radioactivity of building materials. Office for Official Publications of the European Communities, Luxembourg, pp. 98.

ICRP: International Commission on Radiological Protection, 2003. ICRP Publication 90: Biological Effects after Prenatal Irradiation (Embryo and Fetus)Annals of the ICRP, Vol. 33 (1-2), Oxford.

Kendall, G., et al., 2009. Dose to red bone marrow of infants, children and adults from radiation of natural origin. J. Radiol. Prot. 29, 123–138.

Kendall, G.M., et al., 2013. A record-based case-control study of natural background radiation and the incidence of childhood leukaemia and other cancers in Great Britain during 1980–2006. Leukemia 27, 3–9.

Kendall, G.M., et al., 2016. Levels of naturally occurring gamma radiation measured in British homes and their prediction in particular residences. Radiat. Environ. Biophys. 55, 103–124.

Kendall, G.M., et al., 2015. Residential mobility and associated factors in relation to the assessment of exposure to naturally occurring radiation in studies of childhood cancer. J. Radiol. Prot. 35, 835–868.

Kinlen, L., 2012. An examination, with a meta-analysis, of studies of childhood leukaemia in relation to population mixing. Br. J. Cancer 107, 1163–1168.

Kreis, C., et al., 2016. Space-time clustering of childhood cancers in Switzerland: a nationwide study. Int. J. Cancer 138, 2127–2135.

Kroll, M.E., et al., 2010. Childhood cancer and magnetic fields from high-voltage power lines in England and Wales: a case-control study. Br. J. Cancer 103, 1122–1127.

Little, M.P., et al., 2010. The statistical power of epidemiological studies analyzing the relationship between exposure to ionizing radiation and cancer, with special reference to childhood leukemia and natural background radiation. Radiat. Res. 174, 387–402.

Mäkeläinen, I., et al., 2005. Indoor occupancy and radon exposure in Finland. Radioact. Environ. 7, 687–693.

Miles, J.C.H., Appleton, J.D., 2005. Mapping variation in radon potential both between and within geological units. J. Radiol. Prot. 25, 257–276.

National Research Council (NRC), 2006. Committee to Assess Health Risks From Exposure to Low Levels of Ionising Radiation. Health Risks From Exposure to Low Levels of Ionizing Radiation: BEIR VII Phase2. National Academies Press, Washington, DC.

Nikkilä, A., et al., 2016. Background radiation and childhood leukemia: a nationwide register-based case-control study. Int. J. Cancer 139, 1975–1982.

Raaschou-Nielsen, O., et al., 2008. Domestic radon and childhood cancer in Denmark. Epidemiology 19, 536–543.

Raaschou-Nielsen, O., et al., 2001. Air pollution from traffic at the residence of children with cancer. Am. J. Epidemiol. 153, 433–443.

Spix, C., et al., 2017. Background gamma radiation and childhood cancer in Germany: an ecological study. Radiat. Environ. Biophys. 56, 127–138.

Spycher, B.D., et al., 2015. Background ionizing radiation and the risk of childhood cancer: a census-based nationwide cohort study. Environ. Health Perspect. 123, 622–628.

UK Childhood Cancer Study Investigators, 1999. Exposure to power-frequency magnetic fields and the risk of childhood cancer. Lancet. 354, 1925–1931.

UK Childhood Cancer Study Investigators, 2002a. The United Kingdom childhood cancer study of exposure to domestic sources of ionising radiation: 1: radon gas. Br. J. Cancer 86, 1721–1726.

UK Childhood Cancer Study Investigators, 2002b. The United Kingdom Childhood Cancer Study of exposure to domestic sources of ionising radiation: 2: gamma radiation. Br. J. Cancer 86, 1727–1731.
UNSCEAR, 2000. United Nations Scientific Committee on the Effects of Atomic Radiation Sources and effects of ionising radiation 2000 Report to the General Assembly with Scientific Annexes. United Nations, New York, 2000.
UNSCEAR, 2008a. United Nations Scientific Committee on the Effects of Atomic Radiation 2006 Report. Annex A. Epidemiological studies of radiation and cancer. New York, pp. 13–322.
UNSCEAR, 2008b. United Nations Scientific Committee on the Effects of Atomic Radiation Report to the General Assembly, Volume 1: Effects of Ionising Radiation with Scientific Annexes A and B. United Nations, New York.
UNSCEAR, 2013. United Nations Scientific Committee on the Effects of Atomic Radiation 2013 Report Sources, Effects and Risks of Ionising radiation.
Warnery, E., et al., 2015. Indoor terrestrial gamma dose rate mapping in France: a case study using two different geostatistical models. J. Environ. Radioact. 139, 140–148.

PUBLICATION III

Radiation exposure from computerized tomography and risk of childhood leukemia: Finnish register-based case-control study of childhood leukemia (FRECCLE)

Atte Nikkilä, Jani Raitanen, Olli Lohi, Anssi Auvinen

Haematologica 103(11), 1873-1880
https://doi.org/10.3324/haematol.2018.187716

ARTICLE

Radiation exposure from computerized tomography and risk of childhood leukemia: Finnish register-based case-control study of childhood leukemia (FRECCLE)

Atte Nikkilä,[1] Jani Raitanen,[2,3] Olli Lohi[1,4] and Anssi Auvinen[2,3,5]

[1]Faculty of Medicine and Biosciences, University of Tampere; [2]Faculty of Social Sciences, University of Tampere; [3]UKK Institute for Health Promotion Research, Tampere; [4]Tampere Center for Child Health Research, University of Tampere and Tampere University Hospital and [5]STUK – Radiation and Nuclear Safety Authority, Helsinki, Finland

Haematologica 2018
Volume 103(11):1873-1880

ABSTRACT

The only well-established risk factors for childhood leukemia are high-dose ionizing radiation and Down syndrome. Computerized tomography is a common source of low-dose radiation. In this study, we examined the magnitude of the risk of childhood leukemia after pediatric computed tomography examinations. We evaluated the association of computed tomography scans with risk of childhood leukemia in a nationwide register-based case-control study. Cases (n=1,093) were identified from the population-based Finnish Cancer Registry and three controls, matched by gender and age, were randomly selected for each case from the Population Registry. Information was also obtained on birth weight, maternal smoking, parental socioeconomic status and background gamma radiation. Data on computed tomography scans were collected from the ten largest hospitals in Finland, covering approximately 87% of all pediatric computed tomography scans. Red bone marrow doses were estimated with NCICT dose calculation software. The data were analyzed using exact conditional logistic regression analysis. A total of 15 cases (1.4%) and ten controls (0.3%) had undergone one or more computed tomography scans, excluding a 2-year latency period. For one or more computed tomography scans, we observed an odds ratio of 2.82 (95% confidence interval: 1.05 – 7.56). Cumulative red bone marrow dose from computed tomography scans showed an excess odds ratio of 0.13 (95% confidence interval: 0.02 – 0.26) per mGy. Our results are consistent with the notion that even low doses of ionizing radiation observably increase the risk of childhood leukemia. However, the observed risk estimates are somewhat higher than those in earlier studies, probably due to random error, although unknown predisposing factors cannot be ruled out.

Correspondence:
atte.nikkila@uta.fi

Received: January 6 2018.
Accepted: June 26, 2018.
Pre-published: July 5, 2018.

doi:10.3324/haematol.2018.187716

Check the online version for the most updated information on this article, online supplements, and information on authorship & disclosures: www.haematologica.org/content/103/11/1873

Introduction

Leukemia is the most common childhood malignancy.[1] The incidence rates of childhood leukemia in Finland are comparable to those in other European countries and show a slight increasing trend up to the 1990s.[2] Acute lymphoblastic leukemia accounts for approximately 85% of all childhood leukemias. The major histological subtype of acute lymphoblastic leukemia is precursor B-cell acute lymphoblastic leukemia (~85%).[1]

Well-established risk factors for childhood leukemia include high doses of ionizing radiation, alkylating chemotherapy agents, as well as Down syndrome and some rare congenital syndromes such as Fanconi anemia, Bloom syndrome and ataxia telangiectasia.[1,3-6] A number of genetic variants have also been associated

with increased risk of leukemia.[7,8] Furthermore, there is reasonably consistent evidence of a slightly increased risk associated with large birth weight relative to gestational time.[9] A higher risk has also been suggested for older parental age, delivery by Cesarean section, and paternal smoking.[10–13] However, daycare attendance, allergic diseases, maternal folic acid supplementation before birth, and early immune stimulation have been suggested to reduce the risk of leukemia.[14–17]

Although high doses of ionizing radiation increase the risk of childhood leukemia, the magnitude of any effect from low doses remains uncertain. Some studies have suggested increased risks associated with background radiation and following x-ray examinations *in utero* and post-natally.[18–22] Computed tomography (CT) imaging has been used for almost four decades and its frequency of utilization increased greatly during the 1980s-1990s. The annual number of scans peaked around year 2002; more recently CT scans have been partly replaced by magnetic resonance imaging in pediatric imaging, partly because of the risk of cancer from ionizing radiation.[23] In 2015, 5,311 pediatric CT scans were performed in the Finnish population of 1,024,000 children under 17 years old, which is a low rate compared to that in many other countries.[23,24]

Four high-quality studies have investigated the association of pediatric CT scans and childhood leukemia.[25–28] The interpretation of the findings must include an evaluation of confounding by indication, i.e. underlying conditions predisposing children to both CT scans and leukemia.[28–30] Nevertheless, the evidence is still limited and the magnitude of the risk needs to be characterized further.

In this study, we examined the magnitude of the risk of childhood leukemia after pediatric CT examinations using a nationwide case-control design with efforts to avoid reverse causation.

Methods

We used a register-based, case-control study with individually matched controls. The key characteristics of the material have been presented previously.[10] Briefly, all cases of childhood leukemia (M9800–M9948 in ICD-O-3) diagnosed in Finland during 1990–2011 (n=1,100) before the age of 15 years were identified from the Finnish Cancer Registry (Figure 1). Three controls were individually matched, by sex and year of birth, for each case from the Population Register Center. In all analyses, a 2-

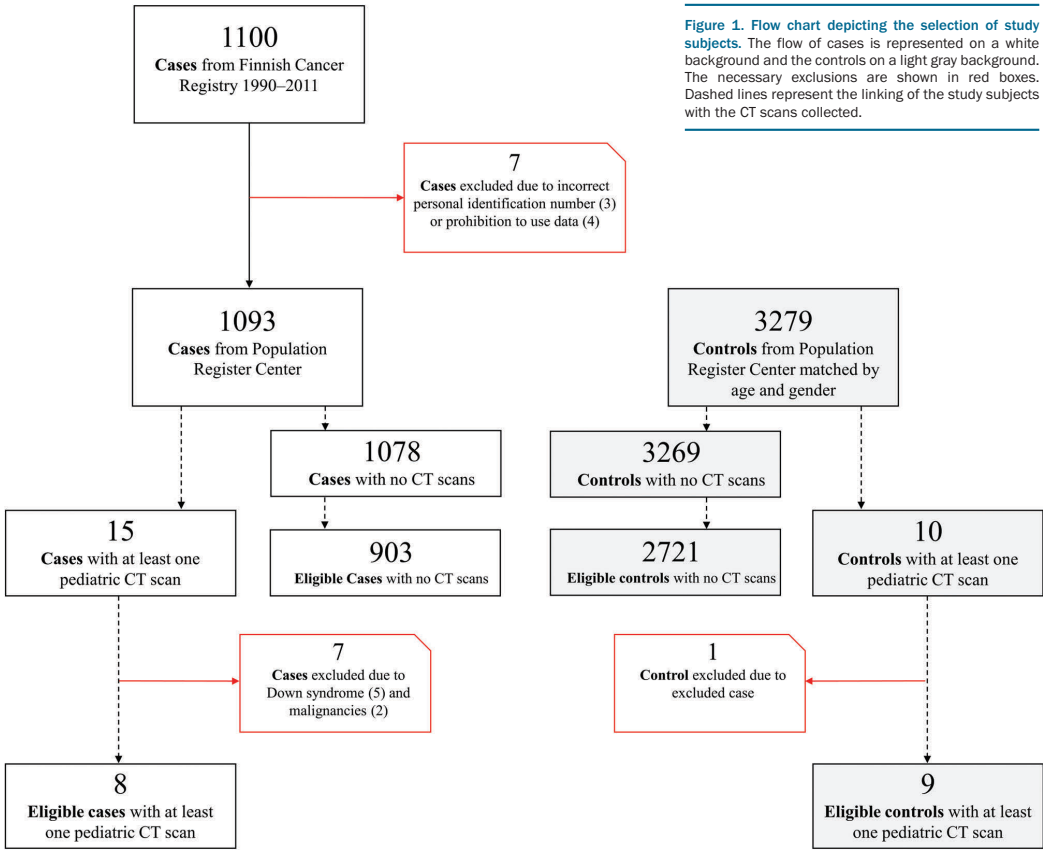


Figure 1. Flow chart depicting the selection of study subjects. The flow of cases is represented on a white background and the controls on a light gray background. The necessary exclusions are shown in red boxes. Dashed lines represent the linking of the study subjects with the CT scans collected.

year latency period was used, in part to deal with reverse causation due to confounding by indication.[31] Also, multiple predisposing factors (*Online Supplementary Table S1*) were accounted for with outpatient register data. The methods are described in more detail in the *Online Supplementary Material*.

We obtained data on all CT scans performed on pediatric patients (<15 years) from all five university hospitals and the five largest central hospitals in Finland (Table 1, Figure 2). The period of data availability varied between hospitals, because radiological databases with information on each CT scan for individual patients were introduced at different times. We estimated that the data from the study hospitals covered 87% of all pediatric CT scans performed in Finland during 1975–2011 (see the *Online Supplementary Material* for details). For each CT scan, we

Table 1. The collection and availability of electronically stored computed tomography scans.

Hospital	City	Data availability	Number of CT scans
Helsinki University Hospital	Helsinki	1990–2011	31,825
Tampere University Hospital	Tampere	1978–2011	9,236
Oulu University Hospital	Oulu	1993–2011	7,513
Turku University Hospital	Turku	1996–2011	7,360
Kuopio University Hospital	Kuopio	1996–2011	5,408
Central Finland Central Hospital	Jyväskylä	2002–2011	2,571
Satakunta Central Hospital	Pori	1995–2011	1,948
Seinäjoki Central Hospital	Seinäjoki	1999–2011	1,759
Päijänne Tavastia Central Hospital	Lahti	2000–2011	1,597
North Karelia Central Hospital	Joensuu	1993–2011	1,191
TOTAL			**80,783**

All Finnish university hospitals are listed first followed by the five chosen central hospitals.

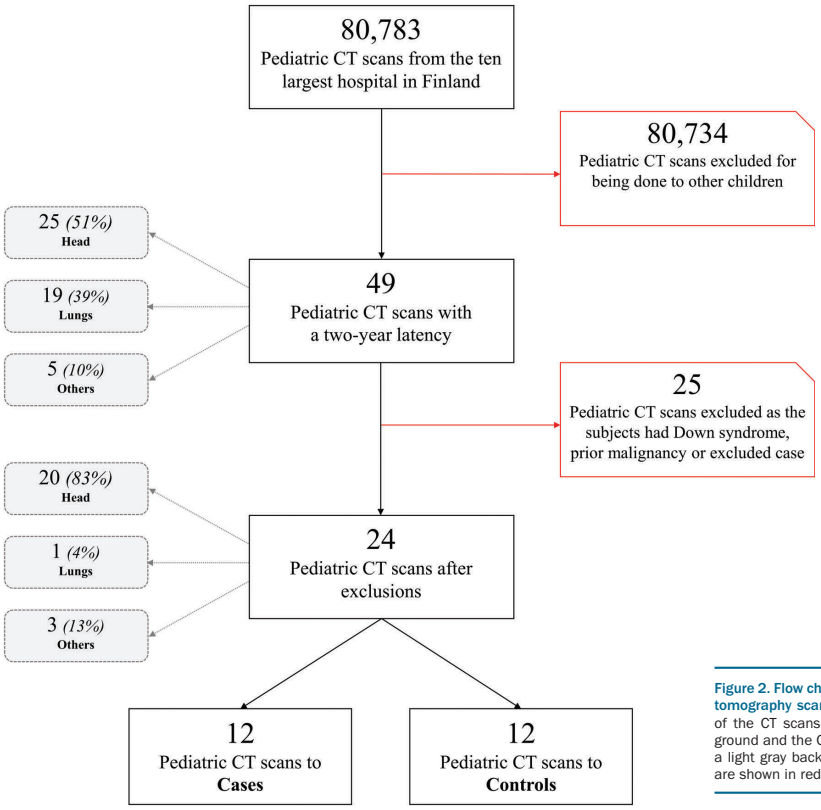


Figure 2. Flow chart linking the collected computed tomography scans to the study subjects The flow of the CT scans is represented on a white background and the CT scans to different body parts on a light gray background. The necessary exclusions are shown in red boxes.

obtained the parameters used for dose assessment including year, body part, use of contrast medium and the number of sequences. Manufacturers and models of CT scanners in each hospital were acquired from the Radiation and Nuclear Safety Authority (STUK). For dose calculations, we assumed in the main analysis that each CT scan was performed using the latest CT scanner available at the hospital.

Data on a total of 80,783 pediatric CT scans were obtained and of those, 49 CT scans were performed on the study subjects, excluding the 2-year latency period (Table 1). Half (n=25) were head scans, and 19 were lung scans. Of the CT scans, 36 were performed on 15 (1.4%) cases and 13 scans on 10 (0.3%) controls.

The CT scan parameters were obtained based on expert opinion of an experienced hospital physicist (*Online Supplementary Table S2*). The doses were estimated using the NCICT software (v1.2).[32] Age- and sex-specific pediatric software phantoms (for neonates, and children aged 1, 5, 10, and 15 years) were used. The input for dose calculation also included the scanner manufacturer and model. If data were available only on the manufacturer, a manufacturer-specific average was used. It was assumed that a head or body filter was used, based on the target body part. The cumulative absorbed red bone marrow (RBM) doses were obtained as the sums of absorbed RBM doses from all CT scans for each study subject. The dose from a scan was multiplied by 1.5 if contrast medium was used, consistent with tissue-specific coefficients suggested for other tissues.[33] Alternative dose estimates were obtained based on values reported in the literature.[34]

We identified subjects with Down syndrome (40 cases and 2 controls) from the Congenital Malformation Register and Care Register for Health Care, and those with a previous malignancy (2 cases) from the Finnish Cancer Registry. They were excluded to avoid confounding by indication (reverse causation). We also collected information on birth weight (large for gestational age) and maternal smoking during pregnancy from the Medical Birth Register, as well as socioeconomic status and education of the parents from Statistics Finland. Residential exposure to background gamma radiation, including natural terrestrial radiation and Chernobyl fallout, was estimated as described previously.[8]

Due to small frequencies, exact conditional logistic regression in SAS 9.4 was used for estimating odds ratios (OR), excess odds ratios and their confidence intervals (CI).[35] Statistical power calculations indicated that the sample size is sufficient for detecting a linear dose-response with an OR of 1.05 or greater per 1 mGy increase in cumulative RBM dose with a statistical power of 80% using asymptotic conditional logistic regression.[36]

The ethical committee of Pirkanmaa Hospital district reviewed the study protocol (tracking number R14074) and, in accordance with Finnish regulations, no informed consent was required for this register-based study. In addition, each hospital approved our study protocol before delivering the data on CT scans. We obtained permission to use data from the Finnish Cancer Registry, the Medical Birth Register, Care Register for Health Care and Congenital Malformation Register from the National Institute of Health and Welfare (1774/5.05.00/2014), as well as census data on socioeconomic status from Statistics Finland (TK-52-306-16).

Results

In our nationwide register-based study, after excluding cases with an incorrect personal identification number or prohibition to use their data, we identified 1,093 cases of childhood leukemia diagnosed in 1990-2011. Most of the cases were acute lymphoblastic leukemia (81.1%) or acute myeloid leukemia (13.0%). The median age at diagnosis among cases was 4.52 years (interquartile range, IQR 2.72 – 8.23). Of the cases and controls, 52% were male (Table 2). The criteria for large for gestational age were met by 121 (13.3%) of the cases and 275 (9.9%) of the controls.

Table 2. The characteristics of cases and controls before any exclusions.

	Cases (n=1,093)	Controls (n=3,279)	*P*
Gender			
Female	48.0% (525)	48.0% (1575)	
Male	52.0% (568)	52.0% (1704)	
Large for gestational age			
No	86.7% (788)	90.1% (2493)	
Ys	13.3% (121)	9.9% (275)	0.001
Missing	184	511	
Mother's smoked during pregnancy			
No	83.1% (742)	84.5% (2296)	
Yes	16.9% (151)	15.5% (420)	0.096
Missing	200	563	
Down syndrome			
No	96.3% (1053)	99.9% (3277)	
Yes	3.7% (40)	0.1% (2)	<0.001
Parents' education			
Mother			
Upper secondary	48.5% (530)	50.6% (1659)	ref
Bachelor's degree	22.3% (244)	23.1% (756)	0.869
Master's or doctor's degree	10.2% (112)	9.8% (321)	0.406
Missing	18.9% (207)	16.6% (543)	
Father			
Upper secondary	52.0% (568)	51.4% (1685)	ref
Bachelor's degree	15.2% (166)	16.2% (532)	0.423
Master's or doctor's degree	10.0% (110)	10.2% (334)	0.880
Missing	22.8% (249)	22.2% (728)	
Parents' socioeconomic status			
Mother			
Self-employed	7.7% (84)	8.3% (273)	ref
Upper level employee	16.1% (176)	15.7% (514)	0.477
Lower level employee	34.8% (380)	34.5% (1130)	0.521
Manual worker	21.4% (231)	20.6% (674)	0.490
Other	18.2% (199)	20.3% (664)	0.859
Missing	2.1% (23)	0.7% (24)	
Father			
Self-employed	13.9% (152)	12.0% (395)	ref
Upper level employee	17.6% (192)	18.2% (596)	0.178
Lower level employee	18.3% (197)	17.9% (587)	0.273
Manual worker	34.0% (372)	35.0% (1148)	0.170
Other	12.4% (135)	14.3% (469)	0.036
Missing	4.1% (45)	2.6% (84)	
Age at leukemia diagnosis, years			
0 – 2	14.3% (156)		
2 – 7	55.5% (605)		
7 – 15	33.4% (332)		
Leukemia type			
Pre-B-ALL	75.6% (826)		
Pre-T-ALL	5.9% (64)		
Unclassified ALL	1.8% (20)		
Acute myeloid leukemia	13.6% (149)		
Other	3.1% (34)		

The reported *P*-values are from an univariate conditional logistic regression model. The non-binary variables were treated as factors and the reference categories are marked with "ref". ALL: acute lymphoblastic leukemia.

After exclusions, eight cases (0.7%) and nine controls (0.3%) had undergone at least one CT scan. The median RBM dose was 10.1 mGy (IQR 4.79 – 13.6) for the exposed cases and 6.29 mGy (IQR 5.69 – 7.14) for the exposed controls (Figure 3). The corresponding literature-based values were 26.5 mGy and 17.6 mGy. The RBM doses calculated with NCICT from thoracic CT scans varied between 1.8 and 6.8 mGy (median 4.0 mGy) and similarly, the doses for head CT scans varied between 1.6 and 10.7 mGy (median 6.6 mGy).

The OR for any *versus* no CT was 2.82 (95% CI: 1.05 – 7.56). For two or more pediatric CT scans, the OR was 5.22 (95% CI: 0.89 – 69.9). For any head CT scans, an OR of 4.00 (95% CI: 1.39 – 11.5) was obtained.

The overall excess OR of childhood leukemia was 0.13 (95% CI: 0.02 – 0.26) per mGy of absorbed RBM dose calculated with the NCICT software (Table 3). Using the cumulative RBM dose estimates from the literature, an excess OR of 0.05 (95% CI: 0.01 – 0.10) per mGy was obtained. In an analysis by dose tertile calculated with NCICT, the excess OR relative to zero dose were 1.26 (95% CI: -0.50 – 10.1) for the first group, 0.09 (95% CI: -0.89 – 10.5) for the second, and 5.00 (95% CI: 0.10 – 31.7) for the last (Figure 4).

For the most common subtype, precursor B-cell acute lymphoblastic leukemia, the excess OR per mGy was 0.14 (95% CI: 0.02 – 0.29) using estimates from NCICT and 0.06 (0.01 – 0.11) for literature-based estimates. The excess OR for any *versus* no CT scans was 2.25 (95% CI: 0.08 – 8.75) for acute lymphoblastic leukemia and 2.88 (95% CI: 0.22 – 11.4) for precursor B-cell acute lymphoblastic leukemia. In the analysis by age at diagnosis/reference date, the excess OR for any *versus* no CT scans was 3.50 (95% CI: -0.25 – 25.9) for children aged 2 – <7 years and 1.27 (95% CI: -0.32 – 6.54) for those aged 7 – <15 years.

Covariate (confounder) adjustments (large for gestational age, maternal smoking during pregnancy, parental education and parental socioeconomic status) did not alter the OR for CT exposure by more than 0.05 units, with the exception of maternal smoking, which increased the OR related to the number of pediatric CT scans (0 *versus* 1 or more) (approximately 0.10 units). Nevertheless, we preferred the unadjusted model, as missing data on maternal smoking resulted in wider confidence intervals for the main variables.

The OR were higher when the subjects with Down syndrome were not excluded (for 1 or more CT scans OR=5.21, 95% CI: 2.19 – 12.4 and for cumulative RBM dose excess OR=0.19 per mGy, 95% CI: 0.07 – 0.32). No evidence of a different effect of the RBM doses on leukemia risk for subjects with or without Down syn-

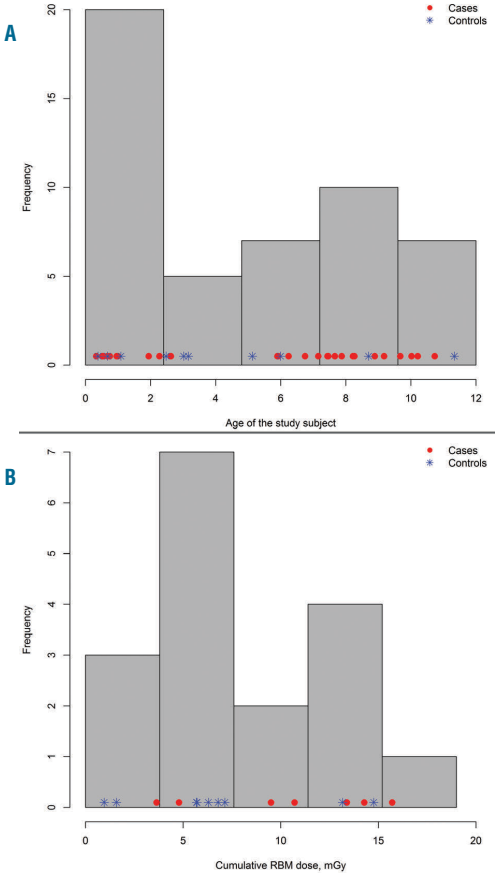


Figure 3. Histograms of (A) the ages of the subjects at the time of computed tomography scan and (B) the cumulative doses received by the subjects, calculated with NCICT.

Table 3. The frequencies of computed tomography scans for subjects >2 years old at the reference date and odds ratios calculated with exact conditional logistic regression.

	Cases	Controls	OR (95% CI)
	911	2730	
CT scans			
0	903	2721	
1	4	7	1.85 (0.39, 7.36)
2 or more	4	2	6.22 (0.89, 68.9)
by type (1 or more)			
ALL	7	7	3.25 (1.08, 9.75)
pre-B-ALL	7	6	3.88 (1.22, 12.4)
by age-group (1 or more)			
2 - <7 years	3	2	4.50 (0.75, 26.9)
7 - <15 years	5	7	2.27 (0.68, 7.54)
by dose index (NCICT/literature)			
low, 4.79[a]/11.6[b] mGy	3	4	2.26 (0.50, 10.1)
medium, 6.72[a]/19. 8[b] mGy	1	3	1.09 (0.11, 10.5)
high, 13.8[a]/33.2[b] mGy	4	2	6.00 (1.10, 32.8)
per 1 mGy (NCICT)			
TOTAL			1.13 (1.02, 1.26)
pre-B-ALL			1.14 (1.02, 1.29)
per 1 mGy (literature)			
TOTAL			1.05 (1.01, 1.10)
pre-B-ALL			1.06 (1.01, 1.11)

The reference group for all calculated odds ratio (OR) is zero CT scans for categorical variables. Study subjects with Down syndrome or cancer diagnoses were excluded. All reported OR are from an unadjusted model. The median doses for dose-index classes calculated with NCICT are marked with [a]. The respective class medians based on literature are marked with [b]. ALL: acute lymphoblastic leukemia; pre-B-ALL: precursor B-cell acute lymphoblastic leukemia.

drome was found to suggest effect modification (interaction *P*=0.99).

When the oldest possible CT scanner (at maximum, 10 years old) at the hospital was used in dose estimation instead of the most modern CT scanner, the median cumulative RBM dose for cases was 9.71 mGy (IQR 7.09 – 18.7) and for controls 7.14 mGy (IQR 5.71 – 12.6), with an excess OR of 0.11 (95% CI: 0.02 – 0.22) per mGy.

When the cumulative RBM dose from terrestrial gamma radiation and Chernobyl fallout was included in the model, the OR for cumulative RBM dose from pediatric CT scans remained unchanged. The median cumulative dose from residential gamma radiation was 1.96 mSv for cases and 1.90 mSv for controls.

The distributions of cities of the last addresses of cases and controls were analyzed to evaluate whether cases and controls belonged to catchment populations of different hospitals, which might have caused differential misclassification due to contrasting data availability. No difference in the distributions was noted (chi-squared test, *P*=0.30). The age and CT scan years of the subjects are reported in *Online Supplementary Table S3*.

Discussion

We estimated the impact of radiation exposure from pediatric CT scans on risk of childhood leukemia in a nationwide register-based case-control study in Finland. Overall, a statistically significant increase in risk per mGy of RBM absorbed dose was found. The central estimate is larger than in previous studies, but the confidence intervals overlap with earlier results, and the effect size is compatible with extrapolation from high-dose studies. The higher main point estimate is likely influenced by random error, as the dose estimates were imprecise due to lack of detail in dosimetric data, including parameter values used for the scanner. It is also possible that the typical values based on expert opinion are representative of current procedures, but may underestimate doses from older examinations, which could inflate the risk estimates per unit dose. However, our site-specific dose estimates calculated with NCICT were quite comparable with those reported in the British study.[25] We minimized the potential for systematic error by adjusting for several confounders and used consistent procedures for the cases and controls. The risk estimates were slightly higher for precursor B-cell acute lymphoblastic leukemia than for other leukemias, but the difference was not statistically significant.

Two large studies have been published on the subject prior to ours. The cohort studies from the United Kingdom and Australia reported a significant risk of childhood leukemia associated with RBM dose from pediatric CT scans.[25,26] Pearce *et al.* found an excess relative risk of 0.04 per mGy and Mathews *et al.* reported a relative risk of 1.2 for one or more CT scans with an excess relative risk of 0.04 per mGy. The Australian cohort had 211 exposed leukemia cases and the UK study 74. A smaller German cohort study reported an increased leukemia incidence following two or more CT scans, but a non-significant dose-response based on 12 exposed cases.[27] Based on the Life Span Study in Japan, the extrapolated excess relative risk for childhood exposure would be approximately 0.05 per mGy.[37]

Other major sources of ionizing radiation were taken into consideration by including cumulative RBM doses from terrestrial gamma radiation and Chernobyl fallout, and this did not affect the results. In our data, the average cumulative RBM dose from CT for the controls was only 0.002 mGy, which is approximately 0.1% of the average annual RBM dose in Finland.[38] We accounted for medical use of radiation, to which tomography scans make the largest contribution, and terrestrial gamma radiation, which accounts for nearly two-thirds of average annual radiation to the RBM in Finland.[23,39] In addition, there is little evidence to assume that other sources of ionizing radiation, such as cosmic radiation or internal exposure to natural radioisotopes, would distribute unequally among the cases and controls.

The coefficient 1.5 for incremental dose due to CT imaging with contrast medium was chosen pragmatically based on the coefficients for other body parts, as the effects on RBM dose were not reported separately.[33]

Based on limited population statistics available from the Radiation and Nuclear Safety Authority,[23] roughly 30 CT scans were expected for the controls. However, only 13 scans were recorded among them. This might partly reflect incomplete availability of data, but the estimate of the expected numbers is highly uncertain because of lack of data on pediatric CT scans prior to 2008. It is also worth noting that pediatric CT scans are performed less frequently in Finland than in several other countries.[24]

Our material consists of a comprehensive set of childhood leukemia cases and representative controls, which should eliminate selection bias by virtue of a register-based approach, which required no consent or information from the study subjects or their families. The study period covers the years in which the use of pediatric CT scans was most frequent, as the annual number of pediatric CT scans has been decreasing in Finland since the year 2000.[23] The data on CT scans were obtained from

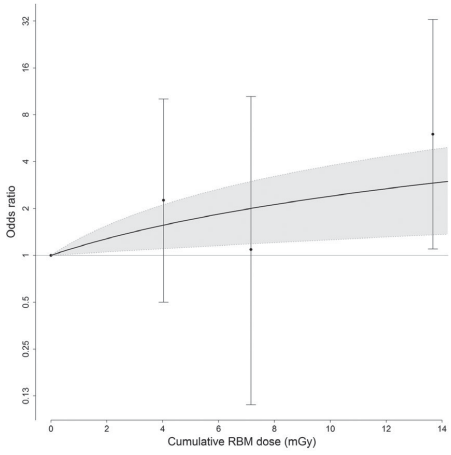


Figure 4. Dose-response curve of cumulative red bone marrow dose from pediatric computed tomography scans and childhood leukemia. The point estimates with 95% confidence intervals are for the three dose index levels and the fitted curve is for the cumulative RBM dose calculated with NCICT. The shaded area represents the 95% confidence interval for the continuous dose-response. The vertical axis is on a binary logarithm scale.

hospital databases to avoid recall bias, and also included the scanner model and use of contrast medium. As in other studies, the most common single CT scan in our analysis was a head scan.[23]

Radiation doses to RBM from the CT scans were calculated using the best available methods, employing NCICT software, with age- and sex-specific phantoms and taking into account the scanner model. The scanning parameters entered into the software were based on the settings and procedures commonly used in Finland, although data were not available for each scan. We also evaluated the effects of choosing the most modern CT scanner at each imaging site and the OR showed robust behavior.

We also had data on several important risk factors including Down syndrome, parental socioeconomic status, large for gestational age and maternal smoking. We were able to incorporate data on cancer predisposing factors, which have been shown to be of importance recently.[28,30] Inclusion of cases with Down syndrome would have increased the risk estimates, possibly because Down syndrome is associated with increased risks of both leukemia and infections.[4,40] We also explored the joint effect of Down syndrome and cumulative RBM dose and found no interaction. Subgroup analyses of exploratory nature were carried out by subtype of childhood leukemia and age at diagnosis, although these were underpowered.

Our study has some shortcomings. We were able to obtain data from all ten hospitals only after 2002, thus exposure assessment is not uniformly complete for subjects born prior to that year. Only a minor improvement in statistical power would have been reached by collecting pediatric CT scans from the rest of the imaging centers in Finland. In addition, there is no reason to assume that the missed CT scans would have been unequally distributed for the cases and controls, i.e. result in differential misclassification. For dose estimation, complete information on the scanning parameters is included in the modern picture archiving systems, but was not available before the year 2000. Use of parameters for each individual scan would have provided more accurate dose estimates. The unexpectedly lower median dose of cases for older scanners found in our sensitivity analysis may be due to random error. The number of different CT scanners in our analysis was limited and thus the estimates of average dose were imprecise.

Our results support the notion that even small doses of radiation from pediatric CT scans produce a small, but detectable increase in leukemia risk. In the subgroup analyses, we observed no substantial differences by age or leukemia subtype, although slightly higher risks were found for precursor B-cell acute lymphoblastic leukemia.

Acknowledgments

The authors would like to thank Isabelle Thierry-Chef (IARC), Hannele Niiniviita (Turku University Hospital), Juha Suutari (STUK) and Rebecca Smith-Bindman (University of California, San Francisco) for their valuable suggestions and comments regarding modeling of CT scan doses with the data available in Finland. We are also grateful to Dr Choonsik Lee (National Cancer Institute) for his insightful comments related to modeling contrast media and for providing us with his state-of-the-art dose calculation software (NCICT) and Anniriikka Rantala (STUK) for collecting the data on CT scanners used in Finland. Päivi Laarne's (Tampere University Hospital) crucial input regarding the scanning parameters enabled us to use NCICT software. Funding for the study was obtained from the Finnish Foundation for Pediatric Research, Väre Foundation for Pediatric Cancer Research and Competitive State Research Financing of the Expert Responsibility area of Tampere University Hospital (9T030 and 9U030).

References

1. Eden T. Aetiology of childhood leukaemia. Cancer Treat Rev. 2010;36(4):286–297.
2. Cancer Incidence, Mortality, Prevalence and Survival in the Nordic Countries, Version 8.0 (20.12.2017).
3. Hsu W-L, Preston DL, Soda M, et al. The incidence of leukemia, lymphoma and multiple myeloma among atomic bomb survivors: 1950-2001. Radiat Res. 2013;179(3): 361–382.
4. Hasle H. Pattern of malignant disorders in individuals with Down's syndrome. Lancet Oncol. 2001;2(7):429–436.
5. Haddy N, Cécile Le Deley M, Samand A, et al. Role of radiotherapy and chemotherapy in the risk of secondary leukaemia after a solid tumour in childhood. Eur J Cancer. 2006;42(16):2757–2764.
6. Kamiya K, Ozasa K, Akiba S, et al. Long-term effects of radiation exposure on health. Lancet. 2015;386(9992):469–478.
7. Wiemels JL, Walsh KM, de Smith AJ, et al. GWAS in childhood acute lymphoblastic leukemia reveals novel genetic associations at chromosomes 17q12 and 8q24.21. Nat Commun. 2018;9(1):286.
8. Ripperger T, Bielack SS, Borkhardt A, et al. Childhood cancer predisposition syndromes-A concise review and recommendations by the Cancer Predisposition Working Group of the Society for Pediatric Oncology and Hematology. Am J Med Genet A. 2017;173(4):1017–1037.
9. Milne E, Greenop KR, Metayer C, et al. Fetal growth and childhood acute lymphoblastic leukemia: findings from the Childhood Leukemia International Consortium. Int J Cancer. 2013;133(12):2968-2979.
10. Nikkilä A, Erme S, Arvela H, et al. Background radiation and childhood leukemia: a nationwide register-based case-control study. Int J Cancer. 2016;139(9): 1975–1982.
11. Marcotte EL, Thomopoulos TP, Infante-Rivard C, et al. Caesarean delivery and risk of childhood leukaemia: a pooled analysis from the Childhood Leukemia International Consortium (CLIC). Lancet Haematol. 2016;3(4):e176–e185.
12. Kendall GM, Little MP, Wakeford R, et al. A record-based case-control study of natural background radiation and the incidence of childhood leukaemia and other cancers in Great Britain during 1980 – 2006. Leukemia. 2013;27(1):3–9.
13. Metayer C, Petridou E, Aranguré JMM, et al. Parental tobacco smoking and acute myeloid leukemia: the Childhood Leukemia International Consortium. Am J Epidemiol. 2016;184(4):261–273.
14. Metayer C, Milne E, Dockerty JD, et al. Maternal supplementation with folic acid and other vitamins and risk of leukemia in offspring: a Childhood Leukemia International Consortium study. Epidemiology. 2014;25(6):811–822.
15. Rudant J, Lightfoot T, Urayama KY, et al. Childhood acute lymphoblastic leukemia and indicators of early immune stimulation: a Childhood Leukemia International Consortium study. Am J Epidemiol. 2015; 181(8):549–562.
16. Ma X, Buffler PA, Selvin S, et al. Daycare attendance and risk of childhood acute lymphoblastic leukaemia. Br J Cancer. 2002; 86(9):1419–1424.
17. Linabery AM, Jurek AM, Duval S, Ross JA. The association between atopy and childhood/adolescent leukemia: a meta-analysis. Am J Epidemiol. 2010;171(7):749–764.
18. Shu XO, Potter JD, Linet MS, et al. Diagnostic X-rays and ultrasound exposure and risk of childhood acute lymphoblastic leukemia by immunophenotype. Cancer Epidemiol Biomarkers Prev. 2002;11(2):177–185.
19. Meinert R, Kaletsch U, Kaatsch P, Schüz J, Michaelis J. Associations between child-

hood cancer and ionizing radiation: results of a population-based case-control study in Germany. Cancer Epidemiol Biomarkers Prev. 1999;8(9):793–799.
20. Infante-Rivard C, Mathonnet G, Sinnett D. Risk of childhood leukemia associated with diagnostic irradiation and polymorphisms in DNA repair genes. Environ Health Perspect. 2000;108(6):495.
21. Bartley K, Metayer C, Selvin S, Ducore J, Buffler P. Diagnostic X-rays and risk of childhood leukaemia. Int J Epidemiol. 2010;39(6): 1628–1637.
22. Harbron RW, Chapple C-L, O'Sullivan JJ, et al. Cancer incidence among children and young adults who have undergone x-ray guided cardiac catheterization procedures. Eur J Epidemiol. 2018;33(4):393-401.
23. Suutari J. Radiologisten tutkimusten ja toimenpiteiden määrät vuonna 2015. STUK; 2015.
24. UNSCEAR. Sources, Effects and Risks of Ionizing Radiation, UNSCEAR 2017 Report. United Nations.
25. Pearce MS, Salotti JA, Little MP, et al. Radiation exposure from CT scans in childhood and subsequent risk of leukaemia and brain tumours: a retrospective cohort study. Lancet. 2012;380(9840):499–505.
26. Mathews JD, Forsythe AV, Brady Z, et al. Cancer risk in 680,000 people exposed to computed tomography scans in childhood or adolescence: data linkage study of 11 million Australians. BMJ. 2013;346:f2360.
27. Krille L, Dreger S, Schindel R, et al. Risk of cancer incidence before the age of 15 years after exposure to ionising radiation from computed tomography: results from a German cohort study. Radiat Environ Biophys. 2015;54(1):1–12.
28. Journy N, Rehel J-L, Ducou Le Pointe H, et al. Are the studies on cancer risk from CT scans biased by indication? Elements of answer from a large-scale cohort study in France. Br J Cancer. 2015;112(1):185–193.
29. Berrington De Gonzalez A, Salotti JA, Mchugh K, et al. Relationship between paediatric CT scans and subsequent risk of leukaemia and brain tumours: assessment of the impact of underlying conditions. Br J Cancer. 2016;114(4):388-394.
30. Journy N, Roué T, Cardis E, et al. Childhood CT scans and cancer risk: impact of predisposing factors for cancer on the risk estimates. J Radiol Prot. 2016;36(1):N1–N7.
31. ANNEx BScienti. UNSCEAR 2013 Report.
32. Lee C, Kim KP, Bolch WE, Moroz BE, Folio L. NCICT: a computational solution to estimate organ doses for pediatric and adult patients undergoing CT scans. J Radiol Prot. 2015;35(4):891-909.
33. Amato E, Salamone I, Naso S, Bottari A, Gaeta M, Blandino A. Can contrast media increase organ doses in CT examinations? A clinical study. Am J Roentgenol. 2013;200(6): 1288–1293.
34. Kim KP, Berrington de González A, Pearce MS, et al. Development of a database of organ doses for paediatric and young adult CT scans in the United Kingdom. Radiat Prot Dosimetry. 2012;150(4):415–426.
35. Mehta CR, Patel NR. Exact logistic regression: theory and examples. Stat Med. 1995;14(19):2143–2160.
36. Lachin JM. Sample size evaluation for a multiply matched case-control study using the score test from a conditional logistic (discrete Cox PH) regression model. Stat Med. 2008;(27):2509–2523.
37. Wakeford R, Little MP, Kendall GM. Risk of childhood leukemia after low-level exposure to ionizing radiation. Expert Rev Hematol. 2010;3(3):251–254.
38. Muikku M, Arvela H, Järvinen H, et al. Annoskakku 2004 - Suomalaisten keskimääräinen efektiivinen annos. 2005.
39. Muikku M, Bly R, Lahtinen J, et al. Suomalaisten keskimääräinen efektiivinen annos Suomalaisten keskimääräinen efektiivinen annos. 2014.
40. Ram G, Chinen J. Infections and immunodeficiency in Down syndrome. Clin Exp Immunol. 2011;164(1):9–16.

PUBLICATION IV

Predicting residential radon concentrations in Finland: Model development, validation, and application to childhood leukemia

Atte Nikkilä, Hannu Arvela, Juha Mehtonen, Jani Raitanen, Merja Heinäniemi, Olli Lohi, Anssi Auvinen

Submitted for publication